

| Result No. | Query # |       |        | DB | ID     | Description        |
|------------|---------|-------|--------|----|--------|--------------------|
|            | Score   | Match | Length |    |        |                    |
| 1          | 474.5   | 49.9  | 173    | 2  | G85771 | superoxide dismuta |
| 2          | 474.5   | 49.9  | 173    | 2  | JC6004 | superoxide dismuta |
| 3          | 474.5   | 49.9  | 173    | 2  | C90323 | superoxide dismuta |
| 4          | 449     | 47.3  | 173    | 2  | AF0694 | copper-zinc supero |
| 5          | 420     | 44.2  | 174    | 2  | A32582 | superoxide dismuta |
| 6          | 414     | 43.6  | 171    | 2  | A82183 | superoxide dismuta |
| 7          | 414     | 43.6  | 175    | 2  | H90768 | probable copper/z  |
| 8          | 414     | 43.6  | 274    | 2  | F85741 | hypothetical prote |
| 9          | 413.5   | 43.5  | 154    | 1  | A33893 | superoxide dismuta |
| 10         | 413     | 43.5  | 175    | 2  | E90877 | probable copper/z  |
| 11         | 413     | 43.5  | 328    | 2  | E85842 | probable superoxid |
| 12         | 404     | 42.5  | 201    | 2  | A10409 | superoxide dismuta |
| 13         | 391     | 41.2  | 173    | 1  | DSFOCL | superoxide dismuta |
| 14         | 391     | 41.2  | 187    | 1  | B41654 | superoxide dismuta |
| 15         | 387     | 40.7  | 186    | 2  | E81855 | superoxide dismuta |
| 16         | 385     | 40.5  | 186    | 2  | F81088 | superoxide dismuta |
| 17         | 374     | 39.4  | 199    | 2  | JC5718 | superoxide dismuta |
| 18         | 373     | 39.3  | 187    | 1  | A41654 | superoxide dismuta |
| 19         | 257     | 27.1  | 98     | 2  | I39650 | superoxide dismuta |
| 20         | 237.5   | 25.0  | 182    | 2  | B75383 | superoxide dismuta |
| 21         | 235     | 24.7  | 174    | 2  | A35383 | superoxide dismuta |
| 22         | 234     | 24.6  | 462    | 2  | B75617 | superoxide dismuta |
| 23         | 217     | 22.8  | 87     | 2  | I39485 | superoxide dismuta |
| 24         | 213.5   | 22.5  | 182    | 2  | H97067 | superoxide dismuta |
| 25         | 207     | 21.8  | 171    | 2  | F70321 | superoxide dismuta |
| 26         | 206     | 21.7  | 169    | 2  | E70390 | superoxide dismuta |
| 27         | 198.5   | 20.9  | 176    | 2  | JE0097 | superoxide dismuta |
| 28         | 198.5   | 20.9  | 184    | 2  | S40984 | superoxide dismuta |
| 29         | 198.5   | 20.9  | 221    | 2  | JE0098 | superoxide dismuta |







C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Aug-2004  
C;Accession: A33893  
R;Beck, B.L.; Tabatabai, L.B.; Mayfield, J.E.  
Biochemistry 29, 372-376, 1990  
A;Title: A protein isolated from *Brucella abortus* is a Cu-Zn superoxide dismutase.  
A;Reference number: A33893; MUID:90148961; PMID:2105741  
A;Accession: A33893  
A;Molecule type: protein  
A;Residues: 1-154 <BEC>  
A;Cross-references: UNIPROT:P15453  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Superfamily: Superoxide dismutase [Cu-Zn]  
C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
F;48,50,73,128/Binding site: copper (His) #status predicted  
F;55-150/disulfide bonds: #status predicted  
F;147/Active site: Arg #status predicted

Query Match 43.5%; Score 413.5; DB 1; Length 154;  
Best Local Similarity 52.0%; Pred. No. 1e-30;  
Matches 79; Conservative 19; Mismatches 53; Indels 1; Gaps 1;  
QY 28 SVTSEVHMIDDNGIKQSIGVTFTDQKGLQIKTDLKLPGAGHGPHIHEGSGCGPAEHD 87  
Db 2 STVVKYREALPTGPGKEVTVVISEAPGLHFVKVNEKLTTPGVHGFVHNEPSCAPGEKD 61

QY 88 GHLTAGLQHGHYDPDKTKHGKPLGNHGKGLDPLRVVKADGIKAKETLLAPRL-TVKEIK 146  
Db 62 GKIVPALAAGGHYDPGNTTHHLGPEGDGEMGDLPLRSANADGKVSITVAVPHLKLAEIK 121

QY 147 GRTVMIHAGGDNYSKPLPLGGGGARIAGCVI 178  
Db 122 QRSMLVHVHGDNYSKPEPLGGGGARFACGVI 153

RESULT 10  
E90877  
probable copper/zinc-superoxide dismutase [imported] - *Escherichia coli* (strain O157:H7,  
C;Species: *Escherichia coli*  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 16-Aug-2004  
C;Accession: E90877  
R;Hawashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: E90877  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-175 <HAY>  
A;Cross-references: UNIPROT:Q8X6B6; GB:BA000007; PIDN:BA035412.1; PID:g13361454; GSPDB:G  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs1989  
C;Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 43.5%; Score 413; DB 2; Length 175;  
Best Local Similarity 47.8%; Pred. No. 1.3e-30;  
Matches 88; Conservative 27; Mismatches 53; Indels 16; Gaps 5;  
QY 1 MKIKLFFVTSTVITSLTSSVVLACSVTSE----VHMIDDNGIKQSIGVTFTDQKG 56  
Db 1 MKCKI-----IAAIAMLTA-----ASCGYAAEQEVPVNLVSADGKEVSGIKITIQETPYG 50

QY 57 LOIKTDLKLPGAGHGPHIHEGSGCGPAEHDGHLTAGLQAHGHYDPDKTKHGHEPLG-NG 115  
Db 51 LLFTPALHSLSEGTHGFVHEKGNCPALPKDGPVAAALSAGHFPDKNTGKHLGFWSPDG 110

QY 116 HKGDPLRVVKADGIKAKETLLAPRL-TVKEIKGRTVMIHAGGDNYSKPLPLGGGGARIA 174  
Db 111 HLGDLPALFVTHDGKANYFVPLAPRLNSLKEIKGRSLMLHAGGDNHHDHPEPLGGGGARMA 170  
QY 175 CGVI 178

Db 171 CGII 174

RESULT 11  
E85842  
probable superoxide dismutase Z3312 [imported] - *Escherichia coli* (strain O157:H7, substrain  
C;Species: *Escherichia coli*  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Aug-2004  
C;Accession: E85842  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: E85842  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-328 <STO>  
A;Cross-references: UNIPROT:Q8X6B6; GB:AE005174; NID:g12516362; PIDN:AAG57201.1; GSPDB:G  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: Z3312  
C;Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 43.5%; Score 413; DB 2; Length 328;  
Best Local Similarity 47.8%; Pred. No. 2.8e-30;  
Matches 88; Conservative 27; Mismatches 53; Indels 16; Gaps 5;  
QY 1 MKIKLFFVTSTVITSLTSSVVLACSVTSE----VHMIDDNGIKQSIGVTFTDQKG 56  
Db 154 MKCKI-----IAAIAMLTA-----ASCGYAAEQEVPVNLVSADGKEVSGIKITIQETPYG 203

QY 57 LOIKTDLKLPGAGHGPHIHEGSGCGPAEHDGHLTAGLQAHGHYDPDKTKHGHEPLG-NG 115  
Db 204 LLFTPALHSLSEGTHGFVHEKGNCPALPKDGPVAAALSAGHFPDKNTGKHLGFWSPDG 263

QY 116 HKGDPLRVVKADGIKAKETLLAPRL-TVKEIKGRTVMIHAGGDNYSKPLPLGGGGARIA 174  
Db 264 HLGDLPALFVTHDGKANYFVPLAPRLNSLKEIKGRSLMLHAGGDNHHDHPEPLGGGGARMA 323

QY 175 CGVI 178  
Db 324 CGII 327

RESULT 12  
A10409  
superoxide dismutase (EC 1.15.1.1) precursor [imported] - *Yersinia pestis* (strain CO92)  
C;Species: *Yersinia pestis*  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 16-Aug-2004  
C;Accession: A10409  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, E  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A;Reference number: AB00001; MUID:21470413; PMID:11586360  
A;Accession: A10409  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-201 <KUR>  
A;Cross-references: UNIPROT:Q8ZEN3; GB:AL590842; PIDN:CAC92605.1; PID:g15981301; GSPDB:G  
C;Genetics:  
A;Gene: sodC  
C;Superfamily: Superoxide dismutase [Cu-Zn]  
C;Keywords: oxidoreductase

Query Match 42.5%; Score 404; DB 2; Length 201;  
Best Local Similarity 54.2%; Pred. No. 1e-29;  
Matches 83; Conservative 18; Mismatches 50; Indels 2; Gaps 2;  
QY 28 SVTSEVHMIDDNGIKQSIGVTFTDQKGLQIKTDLKLPGAGHGPHIHEGSGCGPAEHD 87



C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 16-Aug-2004  
C;Accession: F81088  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755; PMID:10710307  
A;Accession: F81088  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-186 <FET>  
A;Cross-references: UNIPROT:Q59623; GB:AB002488; NID:G7226631; PIDN:AAF4176  
A;Experimental source: serogroup B, strain MC58  
C;Genetics:  
A;Gene: NMB1398  
C;Superfamily: Superoxide dismutase [Cu-Zn]  
C;Keywords: metalloprotein; oxidoreductase  
F;179/Active site: Arg #status predicted

Query Match 40.5%; Score 385; DB 2; Length 186;  
Best Local Similarity 50.3%; Pred. No. 5.2e-28;  
Matches 77; Conservative 22; Mismatches 52; Indels 2; Gaps 2;

QY 28 SVTSEVHMDDNGIKQSIGTFTTDTDKGLQIKTDLKLPGAGHGFHIHGGSCGPAEHD 87  
DB 33 SIEVKVQQLDPVNGNKDVGTVITESTNYGLVFPDQLGLSEGLHGFHIHNSCEPKEXE 92

QY 88 GHLTAGLQAHGHYDPDKTKGHEGP-LGNHGKGLPRLVVKADGIAKETLLAPRL-TVKEI 145  
DB 93 GKLTAGLGGGWDPKAGKHQGYPPQDDAHLGDLPALTVLHDGTATNPVLAPRLKHLDDV 152

QY 146 KGRVTMIHAGDNYSDKPLPLGGGGARIACGVI 178  
DB 153 RGHSMIHAGDGHSDHPAPLGGGGPRMACGVI 185

RESULT 17  
JC5718  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - Haemophilus ducreyi  
C;Species: Haemophilus ducreyi  
C;Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: JC5718  
R;Stevens, M.K.; Hassett, D.J.; Radolf, J.D.; Hansen, E.J.  
Gene 183, 35-40, 1996  
A;Title: Cloning and sequencing of the gene encoding the Cu, Zn-superoxide dismutase of H  
A;Reference number: JC5718; MUID:97149276; PMID:8996084  
A;Accession: JC5718  
A;Molecule type: DNA  
A;Residues: 1-199 <STE>  
A;Cross-references: UNIPROT:Q59452; GB:U47664; NID:gl305411; PIDN:AAB41293.1; PID:gl3054  
C;Genetics:  
A;Gene: sodC  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Superfamily: superoxide dismutase [Cu-Zn]  
C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-199/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
F;92.94,117.173/Binding site: copper (His) #status predicted  
F;99-195/Disulfide bonds: #status predicted  
F;179/Active site: Arg #status predicted

Query Match 39.4%; Score 374; DB 2; Length 199;  
Best Local Similarity 52.7%; Pred. No. 5.8e-27;  
Matches 78; Conservative 15; Mismatches 53; Indels 2; Gaps 2;

QY 33 VHMDDNGIKQSIGTFTTDTDKGLQIKTDLKLPGAGHGFHIHGGSCGPAEHDGHLTA 92  
DB 51 VQQLDPQNGNKDVGTVITESTAYGLVFPKLHDLAHLGHLHFKPSCEPKKDGKLV 110

QY 93 GLQAHGHYDPDKTKGHEPLG-NGHKGDLPLRVVKADGIAKETLLAPRL-TVKEIKGRTV 150  
DB 111 GLGAGGHWDPKQKHGYPWSDDAHMGDLPALFWMHDSATTPVLAPRLKKLAEVKGHSL 170

QY 151 MIHAGDNYSDKPLPLGGGGARIACGVI 178  
DB 171 MIHAGDGHSDHPAPLGGGGPRMACGVI 198

RESULT 18  
A41654  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - Haemophilus influenzae  
C;Species: Haemophilus influenzae  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Aug-2004  
C;Accession: A41654  
R;Kroll, J.S.; Langford, P.R.; Loynds, B.M.  
J. Bacteriol. 173, 7449-7457, 1991  
A;Title: Copper-zinc superoxide dismutase of Haemophilus influenzae and Haemophilus para  
A;Reference number: A41654; MUID:92041655; PMID:1938942  
A;Accession: A41654  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-187 <KRO>  
A;Cross-references: UNIPROT:P25841; GB:M84012; NID:gl48881; PIDN:AAA24953.1; PID:gl48882  
C;Genetics:  
A;Gene: sodC  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Superfamily: Superoxide dismutase [Cu-Zn]  
C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
F;1-35/Domain: signal sequence #status predicted <SIG>  
F;36-187/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
F;87-183/Disulfide bonds: #status predicted  
F;105.114,123,126/Binding site: zinc (His, His, Asp) #status predicted  
F;180/Active site: Arg #status predicted

Query Match 39.3%; Score 373; DB 1; Length 187;  
Best Local Similarity 49.7%; Pred. No. 6.6e-27;  
Matches 76; Conservative 22; Mismatches 53; Indels 2; Gaps 2;

QY 28 SVTSEVHMDDNGIKQSIGTFTTDTDKGLQIKTDLKLPGAGHGFHIHGGSCGPAEHD 87  
DB 34 SIEVKVQQLDPANGNKDVGTVITESTNYGLVFTPNLQGLAEGHGFHIHNSCEPKEXD 93

QY 88 GHLTAGLQAHGHYDPDKTKGHEGP-LGNHGKGLPRLVVKADGIAKETLLAPRL-TVKEI 145  
DB 94 GKLTAGLAAAGHWDKSKAKQHCYPPQDDAHLGDLPALTVLHDGTATNPVLAPRLKKLDEV 153

QY 146 KGRVTMIHAGDNYSDKPLPLGGGGARIACGVI 178  
DB 154 RGHSMIHAGDGHSDHPAPLGGGGPRMACGVI 186

RESULT 19  
I39650  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Actinobacillus pleuropneumoniae (fragments)  
C;Species: Actinobacillus pleuropneumoniae  
C;Date: 16-Aug-1996 #sequence\_revision 05-Mar-1999 #text\_change 09-Jul-2004  
C;Accession: I39650; S22815; S22161  
R;Kroll, J.S.; Langford, P.R.; Wilks, K.E.; Keil, A.D.  
Microbiology 141, 2271-2279, 1995  
A;Title: Bacterial [Cu,Zn]-superoxide dismutase: Phylogenetically distinct from the euka  
A;Reference number: I39485; MUID:96118708; PMID:7496539  
A;Accession: I39650  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-87 <KRO>  
A;Cross-references: UNIPROT:P24702; EMBL:X83123; NID:gl019747; PIDN:CAA58204.1; PID:gl01  
R;Loynds, B.M.; Langford, P.R.; Kroll, J.S.  
Nucleic Acids Res. 20, 615, 1992  
A;Title: recF in Actinobacillus pleuropneumoniae.  
A;Reference number: S22813; MUID:92158680; PMID:1741300  
A;Accession: S22815

A:Molecule type: DNA  
A:Residues: 88-98 <LOY>  
A:Cross-references: EMBL:X63626; NID:g38951; PIDN:CAA45174.1; PID:g1333704  
A:Note: This sequence was submitted to the EMBL Data Library, December 1991  
A:Note: neither the complete nucleic acid sequence nor the complete translation are shown  
C:Genetics:  
A:Gene: sodC  
C:Function:  
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C:Superfamily: superoxide dismutase [Cu-Zn]  
C:Keywords: copper; metalloprotein; oxidoreductase; zinc  
F:18/27/36/39/Banding site: zinc (His, His, His, Asp) #status predicted  
F:91/Active site: Arg #status predicted

Query Match 27.1%; Score 257; DB 2; Length 98;  
Best Local Similarity 56.1%; Pred. No. 1.3e-16;  
Matches 55; Conservative 11; Mismatches 28; Indels 4; Gaps 3;

QY 83 PAEHGHLTAGLQAHGHYDPDKTKGHEGLG-NGHKGDLPRLVVKADGIKATKTLAPRL- 140  
Db 2 PKEGDKLVAGLGCAGGHWPKEKQHGYPWSDNAHLGDLPALFVEHDSATNFVLAPRLK 61

QY 141 TVKEIKGRITVMTHAGDNYSDKPLPLGGGARIACGVI 178  
Db 62 KLDEVKGHSIMTHEGGDNHSDHAPL--GGPRMACGVI 97

RESULT 20  
B75383  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [similarity] - Deinococcus radiodurans (strain  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 16-Aug-2004  
C:Accession: B75383  
R:White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: B75383  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-182 <WHI>  
A:Cross-references: UNIPROT:Q9RU48; GB:AE001998; GB:AE000513; NID:g6459302; PIDN:AAFI110  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1546  
A:Map position: 1  
C:Superfamily: Superoxide dismutase [Cu-Zn]  
C:Keywords: metalloprotein; oxidoreductase

Query Match 25.0%; Score 237.5; DB 2; Length 182;  
Best Local Similarity 32.2%; Pred. No. 1.7e-14;  
Matches 57; Conservative 29; Mismatches 80; Indels 11; Gaps 5;

QY 8 VTSIVTSLTSITSVVLACSVTSEVHMIDNGIKOSIGTIVTFTDQKGLQIKTDLKGLP 67  
Db 5 LTVVPLLAGLGSACADLGOPTVRADLLDQTG--KVTGATTFSPGIGTRVSIKGLK 62

QY 68 AGEHGFHIEGGSGCPA-EHDGHLTAGLQAHGHYDPDKTKGHEGLG---NGHKGDLPRL 123  
Db 63 AGPHGLHIHENPNCNPGPDAGQOTIPFGAAGHFEDEGASHNHGPHARNDDQGHGGLPMI 122

QY 124 VVKADGIKATKTLAPRLTV---KEIKGRITVMTHAGDNYSDKPLPLGGGARIACGVI 177  
Db 123 TVGEDGKRLNFTNRLKMTGTGTGLGRSIVIHADDDYQTN--PAGNSGGRRCGV 177

RESULT 21  
A35383  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 31-Aug-1990 #sequence\_revision 31-Aug-1990 #text\_change 16-Aug-2004

C:Accession: A35383; B87445  
R:Steinman, H.M.; Ely, B.  
J. Bacteriol. 172, 2901-2910, 1990  
A:Title: Copper-zinc superoxide dismutase of Caulobacter crescentus: cloning, sequencing,  
A:Reference number: A35383; MUID:90264275; PMID:2345128  
A:Accession: A35383  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-174 <STE>  
A:Cross-references: UNIPROT:P20379; GB:M55259; NID:gl44282; PIDN:AAA23054.1; PID:gl44283;  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolona  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: B87445  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-174 <STO>  
A:Cross-references: GB:AE005673; NID:gl3422970; PIDN:AAK23558.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CCL579  
C:Function:  
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C:Superfamily: Superoxide dismutase [Cu-Zn]  
C:Keywords: metalloprotein; oxidoreductase  
F:167/Active site: Arg #status predicted

Query Match 24.7%; Score 235; DB 2; Length 174;  
Best Local Similarity 33.9%; Pred. No. 2.8e-14;  
Matches 59; Conservative 23; Mismatches 54; Indels 38; Gaps 6;

QY 18 TSITSVVLACSVTSEVHMIDNGIKOSIGTIVTFTDQKGLQIKTDLKGLPAGEHGFHIE 77  
Db 25 TSAIVVKA-----CGDKAGAVTTEAPHGVLLKLEKGLTPGWHAHFHE 71

QY 78 GSGCGPAEH---DGHL-TAGLQAHGHYDPDKTKGHEGLNGHKGDLPLRLVVKADGIK 133  
Db 72 KGDCGTPDFKSGAGAHVHTAATTVHGLNPD-----NDSGDLNPFNFAADGAATA 121

QY 134 TLLAPRLTVKEIKR-----TVMIHAGDNYSDKPLPLGGGARIACGVI 178  
Db 122 EITYSLPLVSLKAGGCRPALLDAGSSIVVHANPDH--KTQPIGGAGARVACGVI 173

RESULT 22  
B75617  
superoxide dismutase (sodC), Cu-Zn family - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: B75617  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: B75617  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <WHI>  
A:Cross-references: UNIPROT:Q9RYV4; GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAFI1217  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRA0202  
A:Map position: 2  
C:Superfamily: superoxide dismutase (Cu-Zn) with an uncharacterized C-terminal domain

Query Match 24.6%; Score 234; DB 2; Length 462;  
Best Local Similarity 34.4%; Pred. No. 1.1e-13;  
Matches 54; Conservative 24; Mismatches 51; Indels 28; Gaps 5;

```
QY 39 NGIKQSTGTVTTDTDKGLQIKTDLKGLPAGHGFIHHEGSCGPAEHDGHLTAGLQ--- 95
Db 35 DGAGQVVGSARFVQQGAGVQVTVDRGLTPGMHGMVHEFGCTP-----GVDPV 85
QY 96 -----AHGHYDPPDKTGHKEGFL---GNHGKGDLPRLVVKADGIAKETILLAPRLTV-- 142
Db 86 NKVVFGAAGCHDFPMSRNHDTPTQDNKHEGHDTPMLSVGADGVGKASFTSTKISLTG 145
QY 143 -KEIKGRTVMIHAGGDNYSKPLPLGGGGGARIACGVI 178
Db 146 ENGILNRSLVIHANPDY--KTDPAGMSGARERCVI 180

RESULT 23
I39485
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Actinobacillus actinomycetemcomitans (fragm
C;Species: Actinobacillus actinomycetemcomitans
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C;Accession: I39485
R;Kroll, J.S.; Langford, P.R.; Wilks, K.E.; Keil, A.D.
Microbiology 141, 2271-2279, 1995
A;Title: Bacterial [Cu,Zn]-superoxide dismutase: Phylogenetically distinct from the euka
A;Reference number: I39485; MUID:96118708; PMID:7496539
A;Accession: I39485
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-87 <RES>
A;Cross-references: UNIPROT:Q59081; EMBL:X83122; NID:g1019745; PIDN:CAA58203.1; PID:g101
C;Genetics:
A;Gene: sodC
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C;Superfamily: superoxide dismutase [Cu-Zn]
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F18,27,36,39/Binding site: zinc (His, His, Asp) #status predicted

Query Match 22.8%; Score 217; DB 2; Length 87;
Best Local Similarity 52.3%; Pred. No. 5.5e-13;
Matches 45; Conservative 12; Mismatches 27; Indels 2; Gaps 2;

QY 83 PAEHDGHLTAGLQAHGHYDPPDKTGHKEGFLGNGHKGDLPRLVVKADGIAKETILLAPRLT 141
Db 2 PREKDGKLTAGLGAGGHWDPKTKQHYQWQDAHLGDLPLALTVLHDGTANPVLAPRTK 61
QY 142 -VKEIKGRTVMIHAGGDNYSKPLPL 166
Db 62 HLLDDVRGHSMTIHAGGDNHSDHPAPL 87

RESULT 24
H97067
superoxide dismutase, Cu-Zn family [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97067
R;Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: H97067
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-182 <R>
A;Cross-references: UNIPROT:Q97J3; GB:AE001437; PIDN:AAK9331.1; PID:GI5024298; GSPDB:C
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1363
C;Superfamily: superoxide dismutase [Cu-Zn]

Query Match 22.5%; Score 213.5; DB 2; Length 182;
Best Local Similarity 35.1%; Pred. No. 2.7e-12;
Matches 52; Conservative 19; Mismatches 46; Indels 31; Gaps 7;
```

```
QY 46 GTVTETDTDKGLQIKTDLKGLP-----AGEHGFHHEGSCGPAEHDGHLTAGL 94
Db 47 GTVTFTSVPTNGTSSVVVNLGPPYRPMGCRPQVPGPHGFHONGNCTPG-----TAAQ 100
QY 95 QAHGHYDPPDKTGHKEGFL---GNHGKGDLPRLVVKADGIAKETILLAPRLTVKEIKGRTV 150
Db 101 PFPG-----TGEHNPNTQPHGN-FAGDFP-VVFSNGYARMTFTFNKFRFQVIGKSV 152
QY 151 MIHAGGDNYSKPLPLGGGGGARIACGVI 178
Db 153 VLHESPDY--RTQPAGASGRKVACGVI 178

RESULT 25
F70321
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 16-Aug-2004
C;Accession: F70321
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: F70321
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-171 <AQF>
A;Cross-references: UNIPROT:O66602; GB:AE000679; NID:g2982936; PIDN:AAOC6553.1; PID:g298
C;Genetics:
A;Gene: sodC2
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C;Superfamily: Superoxide dismutase [Cu-Zn]
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F164/Active site: Arg #status predicted

Query Match 21.8%; Score 207; DB 2; Length 171;
Best Local Similarity 32.4%; Pred. No. 1e-11;
Matches 59; Conservative 29; Mismatches 68; Indels 26; Gaps 8;

QY 8 VTSVITISLTSITSVLACSVTSEVHMIDONGIKQSTGTVTTFTDCKGLQIKTDLKGLP 67
Db 4 LSGVLGSL--LISASFQDKAHAEINTEG--EVIGKAEIETNSGVLKNAKGLP 59
QY 68 A-GEHGFHHEGSCGPAEHDGHLTAGLQAHGHYDPPDKTGHKEGFLG--NGHKGDLPRLV 124
Db 60 PNAELAFPIHERGECKPPTFK-----SARGHNP--YGKKHGLNPEGPHAGDMPNII 110
QY 125 VKADGIAKETILLAPRLTVKEIK-----GRTVMIHAGGDNYSKPLPLGGGGGARIACG 176
Db 111 TDDKGNVRVQVLPNPFVTLKKGKRNLSFKEGGTALVIHSGPDDY--KSDPAGNAGKRIACG 168
QY 177 VI 178
Db 169 VI 170

RESULT 26
B70390
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 16-Aug-2004
C;Accession: B70390
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: B70390
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
```



[illegible]

A;Molecule type: DNA  
A;Residues: 1-176 <U>  
A;Cross-references: UNIPROT:P34461; DDBJ:AB003324; NID:G3135194; PIDN:BA28262.1; PID:G3135194  
C;Comment: This protein is an extracellular form.  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Superfamily: superoxide dismutase [Cu-Zn]  
C;Keywords: copper; glycoprotein; metalloprotein; oxidoreductase; zinc  
F;56/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;70,72,87,144/Binding site: copper (His) #status predicted  
F;81-170/Disulfide bonds: #status predicted  
F;87,95,104,107/Binding site: zinc (His, His, Asp) #status predicted  
F;167/Active site: Arg #status predicted

d**b** 153 GTSDOSKTTGNAGSRLACGTI 173

|                       |              |                    |                |             |
|-----------------------|--------------|--------------------|----------------|-------------|
| Query Match           | 20.9%;       | Score 198.5;       | DB 2;          | Length 221; |
| Best Local Similarity | 37.6%;       | Pred. No. 8.1e-11; |                |             |
| Matches 53:           | Conservative | 17;                | Mismatches 56; | Indels 15;  |
|                       |              |                    |                | Gaps 4      |



[illegible]

A:Accession: S03608  
A:Molecule type: mRNA  
A:Residues: 1-219 <TEP>  
A:Cross-references: UNIPROT:P10792; EMBL:X14352; NID:g20581; PIDN:CAA32534.1; PID:g20582  
C:Function:  
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C:Superfamily: superoxide dismutase [Cu-Zn]  
C:Keywords: chloroplast; copper; metalloprotein; oxidoreductase; zinc  
F:1-65/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F:66-219/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
F:111,113,128,185/Binding site: copper (His) #status predicted  
F:122-211/Disulfide bonds: #status predicted  
F:128,136,145,148/Binding site: zinc (His, His, His, Asp) #status predicted  
F:208/Active site: Arg #status predicted

Query Match 18.9%; Score 180; DB 2; Length 219;  
Best Local Similarity 29.8%; Pred. No. 4e-09; 68; Indels 30; Gaps 9;  
Matches 54; Conservative 29; Mismatches 29

QY 12 VTISLLTSITSVLWACSVTSEVHMDDNGIKSIGIVTFTDTDKG-LQIKTDLKGLPAGE 70  
DB 50 LTLSSVTSPPRFIVFAATKKAVAL--KGTSNVEGVVTLTQDDGFTVKVITGLAPGL 107  
QY 71 HGFPHIEGGSCPAHDGHLTAGLQAHG-HYDPDKTKGHEGLGN--GHKGDLPRLVVKA 127  
DB 108 HGFPHLEFGD-----TTNGCMSTGPHFNPN--GLTHGAPGDEVRHAGDLGNIFANA 156  
QY 128 DGIAKETLLAPRLTV--KEIKGRVMIH-----AGDNYSDKPLPLGGGGARIACGV 177  
DB 157 SGVAEATLVNDQIPLSGPNSVVGRALVWHEEDDLKGGHELS---LTTGNAGRLIACGV 213  
QY 178 I 178  
DB 214 V 214

RESULT 35

S48021  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 15-Jul-1995 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C:Accession: S48021; S08497  
R:Kardish, N.; Magal, N.; Aviv, D.; Galun, E.  
Plant Mol. Biol. 25, 887-897, 1994  
A:Title: The tomato gene for the chloroplastic Cu-Zn superoxide dismutase: regulation of  
A:Reference number: S48021; MUID:94355661; PMID:8075404  
A:Accession: S48021  
A:Molecule type: DNA  
A:Residues: 1-217 <KAR>  
A:Cross-references: UNIPROT:P14831  
R:Pierl-reves, R.; Nacmias, B.; Aviv, D.; Zeelon, E.P.; Galun, E.  
Plant Mol. Biol. 11, 609-623, 1988  
A:Title: Isolation of two cDNA clones from tomato containing two different superoxide di  
A:Reference number: S08350  
A:Accession: S08497  
A:Molecule type: mRNA  
A:Residues: 1-217 <PER>  
A:Cross-references: EMBL:X14041; NID:g19192; PIDN:CAA32200.1; PID:g19193  
C:Genetics:  
A:Genome: nuclear  
C:Function:  
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C:Superfamily: superoxide dismutase [Cu-Zn]  
C:Keywords: chloroplast; copper; metalloprotein; oxidoreductase; zinc  
F:1-63/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F:64-217/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
F:109,111,126,183/Binding site: copper (His) #status predicted  
F:120-209/Disulfide bonds: #status predicted  
F:126,134,143,146/Binding site: zinc (His, His, His, Asp) #status predicted  
F:206/Active site: Arg #status predicted

Query Match 18.9%; Score 179.5; DB 2; Length 217;  
Best Local Similarity 29.5%; Pred. No. 4.4e-09;

|   |     |   |     |            |     |        |     |      |     |
|---|-----|---|-----|------------|-----|--------|-----|------|-----|
| Matches   | 56; | Conservative  | 26; | Mismatches | 69; | Indels | 39; | Gaps | 10; |
| QY  | 3   | IKLFFVTSIVTISLTSITSVVLACSVTSEVHMIDNGIKQSIGTWTFTDQKG-LQIKT   | 61  |            |     |        |     |      |     |
| Db  | 48  | LTLVAVITPKPLTVFAATKKAVALKGNVSE-----GVVLQDDDGPTTVNV          | 96  |            |     |        |     |      |     |
| QY  | 62  | DLKGLPAGEGHFHIHEGSCGPAEHGHLTAGLQAHG-HYDPDKTKGHEGPLGN--GHKG  | 118 |            |     |        |     |      |     |
| Db  | 97  | RITGLAPGLGHFLHEXGD-----TTNGCMSTGAHFNPKL-THGAP-GDEIRHAG      | 145 |            |     |        |     |      |     |
| QY  | 119 | DLPRVLVKADGIAKETLL---APRLTVKEIKGRTVMIH-----AGGNYSKDPLPLGG   | 168 |            |     |        |     |      |     |
| Db  | 146 | DLGNIVANADGVAEVLVDNQIPLTGPNSVGRALVHLEDDLGKGGHLS---LTTGN     | 202 |            |     |        |     |      |     |
| QY  | 169 | GGARIACGVI  | 178 |            |     |        |     |      |     |
| Db  | 203 | AGRLACGVV   | 212 |            |     |        |     |      |     |
| RESULT 36   |     |   |     |            |     |        |     |      |     |
| S20512  |     |   |     |            |     |        |     |      |     |
| superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Scotch pine (fragment)                       |     |   |     |            |     |        |     |      |     |
| C/Species: Pinus sylvestris (Scotch pine)   |     |   |     |            |     |        |     |      |     |
| C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 09-Jul-2004               |     |   |     |            |     |        |     |      |     |
| C/Accession: S20512; S14613   |     |   |     |            |     |        |     |      |     |
| R/Karpinski, S.; Wingsle, G.; Olsson, O.; Haellgren, J.E.                                 |     |   |     |            |     |        |     |      |     |
| Plant Mol. Biol. 18, 545-555, 1992  |     |   |     |            |     |        |     |      |     |
| A/Title: Characterization of cDNAs encoding CuZn-superoxide dismutases in Scots pine.     |     |   |     |            |     |        |     |      |     |
| A/Reference number: S20511; MUID:92163019; PMID:1371406                                   |     |   |     |            |     |        |     |      |     |
| A/Accession: S20512   |     |   |     |            |     |        |     |      |     |
| A/Molecule type: mRNA   |     |   |     |            |     |        |     |      |     |
| A/Residues: 1-141 <KAR>   |     |   |     |            |     |        |     |      |     |
| A/Cross-references: UNIPROT:P24707; EMBL:X58579; NID:G20696; PIDN:CAA41455.1; PID:G20697  |     |   |     |            |     |        |     |      |     |
| A/Function:   |     |   |     |            |     |        |     |      |     |
| A/Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen   |     |   |     |            |     |        |     |      |     |
| C/Superfamily: superoxide dismutase [Cu-Zn]   |     |   |     |            |     |        |     |      |     |
| C/Keywords: chloroplast; copper; metalloprotein; oxidoreductase; zinc                     |     |   |     |            |     |        |     |      |     |
| F/33,35,50,107/Binding site: copper (His) #status predicted                               |     |   |     |            |     |        |     |      |     |
| F/44-133/Disulfide bonds: #status predicted   |     |   |     |            |     |        |     |      |     |
| F/50,58,67,70/Binding site: zinc (His, His, His, Asp) #status predicted                   |     |   |     |            |     |        |     |      |     |
| F/130/Active site: Arg #status predicted  |     |   |     |            |     |        |     |      |     |
| Query Match 18.8%; Score 179; DB 2; Length 141;   |     |   |     |            |     |        |     |      |     |
| Best Local Similarity 32.9%; Pred. No. 2.9e-09;   |     |   |     |            |     |        |     |      |     |
| Matches 48; Conservative 23; Mismatches 49; Indels 26; Gaps 8;                            |     |   |     |            |     |        |     |      |     |
| QY  | 46  | GVTFITDQKG-LQIKTDLKGLPAGEGHFHIHEGSCGPAEHGHLTAGLQAHG-HYDPD   | 103 |            |     |        |     |      |     |
| Db  | 4   | GVVLQDDDGPTTVKRVITGLTPGKHGFLHEFGD-----TTNGCMSTGSHFNPK       | 54  |            |     |        |     |      |     |
| QY  | 104 | KTGKHEGPLGN-GHKGDPLRVLVKADGIAKETLLAPRLTVK---EIKGRTVMIH----- | 153 |            |     |        |     |      |     |
| Db  | 55  | KL-THGAPEDDVRHAGDLGNIVAGSDGVAEATVDNQIPLSGPDSVIGRALVHLEDDL   | 113 |            |     |        |     |      |     |
| QY  | 154 | -AGGNYSKDPLPLGGGGARIACGVI                                   | 178 |            |     |        |     |      |     |
| Db  | 114 | GKGGHLS---LTTGNAGRLACGVV                                    | 136 |            |     |        |     |      |     |
| RESULT 37   |     |   |     |            |     |        |     |      |     |
| DSWFCZ  |     |   |     |            |     |        |     |      |     |
| superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - swordfish                                    |     |   |     |            |     |        |     |      |     |
| C/Species: Xiphias gladius (swordfish)  |     |   |     |            |     |        |     |      |     |
| C/Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 16-Aug-2004               |     |   |     |            |     |        |     |      |     |
| C/Accession: A00516   |     |   |     |            |     |        |     |      |     |
| R/Rocha, H.A.; Bannister, W.H.; Bannister, J.V.   |     |   |     |            |     |        |     |      |     |
| Eur. J. Biochem. 145, 477-484, 1984   |     |   |     |            |     |        |     |      |     |
| A/Title: The amino-acid sequence of copper/zinc superoxide dismutase from swordfish liver |     |   |     |            |     |        |     |      |     |
| A/Reference number: A00516; MUID:85076642; PMID:6510412                                   |     |   |     |            |     |        |     |      |     |
| A/Accession: A00516   |     |   |     |            |     |        |     |      |     |
| A/Molecule type: protein  |     |   |     |            |     |        |     |      |     |
| A/Residues: 1-151 <ROC>   |     |   |     |            |     |        |     |      |     |
| A/Cross-references: UNIPROT:P03946  |     |   |     |            |     |        |     |      |     |

Db 96 HEYGDITNGICSTGP-----HENPNKL-THGAPDEIRHAGDIGNIVANA 139  
 QY 128 DGAKEKTLT---APRLTVKEIKGRTVMIH-----AGGDNYSDKPLPLGGGCGARIACGV 177  
 Db 140 EGVAEATIVDNQIPLTGPNSVVGRLVHVHQLDQLKGGHLS---LSTGNAGRLACGV 196  
 QY 178 I 178  
 Db 197 V 197

## RESULT 39

S03606  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - fruit fly (*Drosophila virilis*)  
 C;Species: *Drosophila virilis*  
 C;Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 09-Jul-2004  
 C;Accession: S03606  
 R;Kwiatowski, J.; Ayala, F.J., 1989  
 Nucleic Acids Res. 17, 2133, 1989  
 A;Title: *Drosophila virilis* Cu-Zn superoxide dismutase gene sequence.  
 A;Reference number: S03606; MUID:99183628; PMID:2928122  
 A;Accession: S03606  
 A;Molecule type: DNA  
 A;Residues: 1-153 <KWI>  
 A;Cross-references: UNIPROT:P10791; EMBL:X13831; NID:g9204; PIDN:CRA32060.1; PID:g9205  
 C;Genetics:  
 A;Gene: FlyBase:Dvir/Sod  
 A;Cross-references: FlyBase:FBgn0013096  
 A;Introns: 22/3  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: superoxide dismutase [Cu-Zn]  
 C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
 F;45,47,62,119/Binding site: copper (His) #status predicted  
 F;56-145/Disulfide bonds: #status predicted  
 F;62,70,79,82/Binding site: zinc (His, His, Asp) #status predicted  
 F;142/Active site: Arg #status predicted

Query Match 18.6%; Score 176.5; DB 2; Length 153;  
 Best Local Similarity 32.9%; Pred. No. 5.5e-09;  
 Matches 47; Conservative 24; Mismatches 53; Indels 19; Gaps 7;  
 QY 46 GTVTFTDTDKGLQIKT--DLKGLPAGEHGPHHEGSGCPAEHDGHLTAGLQAHGHYDPD 103  
 Db 15 GTVFEQEGECVPKVTGEVTLGAKQGHGFHVEGD---NTNGCMSSG---PHENFY 66  
 QY 104 KTGHEGFLG-NGHKGDPLRLVVKADGIKETLLAPRLTV---KEIKGRTVMIHAGGDNY 159  
 Db 67 QK-EHGAPTENRHLGLGNIANGDGTPVNICDKITLLGANSIIIGRTVVVHADPDDL 125  
 QY 160 S-----DKPLPLGGGCGARIACGV 178  
 Db 126 QKGHLSKTTGNAGARIGCGVI 148

## RESULT 40

S05021  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) A - African clawed frog  
 C;Species: *Xenopus laevis* (African clawed frog)  
 C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
 C;Accession: S05021; S59616; A36699  
 R;Schinina, M.E.; Barra, D.; Bossa, F.; Calabrese, L.; Montesano, L.; Carri, M.T.; Mario  
 Arch. Biochem. Biophys. 272, 507-515, 1989  
 A;Title: Primary structure from amino acid and cDNA sequences of two Cu,Zn superoxide di  
 A;Reference number: S05021; MUID:89321563; PMID:2751312  
 A;Accession: S05021  
 A;Molecule type: mRNA  
 A;Residues: 1-150 <SCH>  
 A;Cross-references: UNIPROT:P13926  
 A;Note: this sequence was confirmed by protein sequencing  
 A;Accession: S59616  
 A;Molecule type: mRNA  
 A;Residues: 1-150 <SCW>

R;Capo, C.R.; Polticelli, F.; Calabrese, L.; Schinina, M.E.; Carri, M.T.; Rotilio, G.  
 Biochem. Biophys. Res. Commun. 173, 1186-1193, 1990  
 A;Title: The Cu,Zn superoxide dismutase isoenzymes of *Xenopus laevis*: purification, ident  
 A;Reference number: A36699; MUID:91097547; PMID:2268321  
 A;Accession: A36699  
 A;Molecule type: protein  
 A;Residues: 1-30 <CAP>  
 A;Note: AA homodimers, BB homodimers, and AB heterodimers were observed; the material seq  
 re A and B differed  
 A;Note: 18-His, 24-Glu, 26-Ala, and 28-Ser (Cu-Zn superoxide dismutase B) were also found  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: superoxide dismutase [Cu-Zn]  
 C;Keywords: copper; heterodimer; homodimer; metalloprotein; oxidoreductase; zinc  
 F;44,46,61,117/Binding site: copper (His) #status predicted  
 F;55-143/Disulfide bonds: #status predicted  
 F;61,69,78,81/Binding site: zinc (His, His, Asp) #status predicted  
 F;140/Active site: Arg #status predicted

Query Match 18.5%; Score 175.5; DB 2; Length 150;  
 Best Local Similarity 33.8%; Pred. No. 6.6e-09;  
 Matches 50; Conservative 18; Mismatches 49; Indels 31; Gaps 7;  
 QY 46 GTVTFTDTDKG-LQIKTDLKLPGAGEHGPHHEGSGCPAEHDGHLTAG---LQAHGHY 100  
 Db 15 GVVRFEQQDDGVTVGKIEGLTDGNHGFHVFVG---NTNGCLSGAGPHFNPQNKHG 70  
 QY 101 DPKTKGHEGFLGNGHKGDPLRLVVKADGIKETLLAPRLTVK---EIKGRTVMIH--- 153  
 Db 71 SPKDADRHVGDLGN-----VTAEGGVAQPKFTDPQISLKGERSIIIGRTAVVHEKQD 121  
 QY 154 ---AGSDNYSDKPLPLGGGCGARIACGV 178  
 Db 122 DLGKGGD---DESLKTGNAGRLACGV 146

Search completed: October 26, 2004, 09:45:37  
 Job time : 38 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 09:28:44 ; Search time 154 Seconds  
(without alignments)

419.294 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 950

Sequence: 1 MKIKLFFVTSIVTISLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_23Sep04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 950   | 100.0       | 180    | 4     | AAB47008 L. intrac |
| 2          | 430   | 45.3        | 213    | 6     | ADA34763 Acinetoba |
| 3          | 414   | 43.6        | 175    | 7     | ADC00431 Enterohae |
| 4          | 413.5 | 43.5        | 154    | 2     | AAR32374 20KD Bruc |
| 5          | 413   | 43.5        | 175    | 7     | ADC00874 Enterohae |
| 6          | 403.5 | 42.5        | 144    | 7     | ABO62994 Klebsiell |
| 7          | 380.5 | 40.1        | 178    | 7     | ADF04181 Bacterial |
| 8          | 207.5 | 21.8        | 179    | 4     | ABB64175 Drosophil |
| 9          | 186.5 | 20.7        | 181    | 4     | ABB67296 Drosophil |
| 10         | 187.5 | 19.7        | 221    | 5     | AAM52486 Superoxid |
| 11         | 185   | 19.5        | 218    | 5     | AAM52483 Superoxid |
| 12         | 179.5 | 18.9        | 217    | 5     | AAM52484 Superoxid |
| 13         | 178   | 18.7        | 195    | 5     | AAM52477 Superoxid |
| 14         | 178   | 18.7        | 202    | 5     | AAM52485 Superoxid |
| 15         | 175.5 | 18.5        | 150    | 5     | AAM52487 Superoxid |
| 16         | 175.5 | 18.5        | 166    | 5     | AAM52497 Superoxid |
| 17         | 175   | 18.4        | 216    | 3     | AAQ05964 Arabidops |
| 18         | 175   | 18.4        | 216    | 3     | AAQ49481 Arabidops |
| 19         | 175   | 18.4        | 216    | 8     | ADN73523 Thale cre |
| 20         | 175   | 18.4        | 230    | 3     | AAQ49480 Arabidops |
| 21         | 175   | 18.4        | 232    | 3     | AAQ05963 Arabidops |
| 22         | 172.5 | 18.2        | 196    | 2     | AAR24225 GAG fusio |
| 23         | 171.5 | 18.1        | 183    | 2     | AAR27948 GAG fusio |
| 24         | 171.5 | 18.1        | 185    | 2     | AAR27938 GAG fusio |
| 25         | 171.5 | 18.1        | 193    | 2     | AAR27947 GAG fusio |

|    |       |      |     |   |          |                    |
|----|-------|------|-----|---|----------|--------------------|
| 26 | 171.5 | 18.1 | 197 | 2 | AAR27937 | Aar27937 GAG fusio |
| 27 | 171.5 | 18.1 | 203 | 2 | AAR24235 | Aar24235 GAG fusio |
| 28 | 171.5 | 18.1 | 209 | 2 | AAR24233 | Aar24233 GAG fusio |
| 29 | 171   | 18.0 | 152 | 3 | AAG11627 | Aag11627 Arabidops |
| 30 | 171   | 18.0 | 152 | 3 | AAG07549 | Aag07549 Arabidops |
| 31 | 171   | 18.0 | 152 | 3 | AAG47395 | Aag47395 Arabidops |
| 32 | 171   | 18.0 | 152 | 3 | AAG31513 | Aag31513 Arabidops |
| 33 | 171   | 18.0 | 152 | 3 | AAG31512 | Aag31512 Arabidops |
| 34 | 170.5 | 17.9 | 174 | 1 | AAP81018 | Aap81018 Sequence  |
| 35 | 170   | 17.9 | 152 | 3 | AAG06932 | Aag06932 Arabidops |
| 36 | 170   | 17.9 | 264 | 4 | ABB59872 | Abb59872 Drosophil |
| 37 | 169.5 | 17.8 | 154 | 2 | AAW17901 | Aaw17901 Human sup |
| 38 | 169.5 | 17.8 | 183 | 2 | AAR27951 | Aar27951 GAG fusio |
| 39 | 169.5 | 17.8 | 184 | 2 | AAR27941 | Aar27941 GAG fusio |
| 40 | 169.5 | 17.8 | 186 | 2 | AAR27946 | Aar27946 GAG fusio |
| 41 | 169.5 | 17.8 | 186 | 2 | AAR27943 | Aar27943 GAG fusio |
| 42 | 169.5 | 17.8 | 189 | 2 | AAR27936 | Aar27936 GAG fusio |
| 43 | 169.5 | 17.8 | 189 | 2 | AAR27933 | Aar27933 GAG fusio |
| 44 | 169.5 | 17.8 | 192 | 2 | AAR24229 | Aar24229 GAG fusio |
| 45 | 169.5 | 17.8 | 192 | 2 | AAR24231 | Aar24231 GAG fusio |

#### ALIGNMENTS

RESULT 1

AAB47008

ID AAB47008 standard; protein; 180 AA.

XX AC AAB47008;

XX DT 22-MAR-2001 (first entry)

XX DE L. intracellularis SodC.

XX KW Porcine proliferative enteropathy; immunogen; SodC; antibody; pig;

XX KW vaccine; intestinal infection; serum; blood lymph node; ileum; caecum;

XX KW small intestine; large intestine; faeces; rectal swab; PPE.

XX OS Lawsonia intracellularis.

XX FH Key Location/Qualifiers

XX FT Peptide 1..42

XX FT /note= "Immunogenic peptide fragment"

XX PN WO200069903-A1.

XX ED 23-NOV-2000.

XX PF 11-MAY-2000; 2000WO-AU000436.

XX PR 13-MAY-1999; 99US-0133989P.

XX PA (PFIZ ) PFIZER PROD INC.

XX PA (PIGR-) PIG RES & DEV CORP.

XX PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

XX PI Ankenbauer RG, Hasse D, Panaccio M, Rosey EL, Wright C;

XX DR WPI; 2001-031924/04.

XX N-ESDB; AAC85254.

XX PT Isolated or recombinant polypeptide for treating porcine and avian species against Lawsonia intracellularis infection, comprises, mimics or cross-reacts with the B or T cell epitope of Lawsonia SodC polypeptide.

XX PS Claim 6; Page 79-80; 85pp; English.

XX CC This sequence represents an immunogenic polypeptide, SodC, which is capable of eliciting the production of antibodies against L.

XX CC intracellularis when administered to an avian or porcine animal. This

XX CC polypeptide can be used in a vaccine composition for the prophylaxis or

XX CC treatment of intestinal infection of an animal by Lawsonia. The DNA





Db 1 MKCKI-----IAIAMLTA-----ASCGYAAEVEPMNLVSDGKVSIGKTIQTFPYG 50  
 QY 57 LQIKTDKGLPAGHGFHIEGSGCPAEBHDGHLTAGLQAHGHYDPDKTKHEGPIG-NG 115  
 Db 51 LLFTPALHSLSEGIHGFVHEKGNCPALKDGPVLAALSAGHFDPNKTKHLPWSPDG 110  
 QY 116 HKGDLPLRVKADGIAKETLLAPRL-TVXKIKGRTVMIHAGGDNYSKPLPLGGGGARIA 174  
 Db 111 HPGDLPALFVTHDGKANYPLVAPRLNSLKEIKGRSLMHAGGDNHHDHPEPLGGGGARVA 170  
 QY 175 CGVI 178  
 Db 171 CGII 174

RESULT 4  
 AAR32374  
 ID AAR32374 standard; protein; 154 AA.  
 XX  
 AC AAR32374;  
 XX  
 XX 24-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 30-JUN-1993 (first entry)  
 XX  
 DE 20kD Brucella abortus copper-zinc superoxide dismutase.  
 XX  
 KW CuZnSOD; bovine brucellosis infection; detection.  
 XX  
 OS Brucella melitensis biovar Abortus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..96  
 FT /note= "CL1"  
 FT Region 1..50  
 FT /note= "N-terminal"  
 FT Region 1..7  
 FT /note= "CB1"  
 FT Region 8..37  
 FT /note= "CB2"  
 FT Region 38..91  
 FT /note= "CB3"  
 FT Region 92..154  
 FT /note= "CB4-CB5? Not clear from specification"  
 FT Region 97..123  
 FT /note= "CL2"  
 FT Region 119..138  
 FT /note= "SA10"  
 FT Region 124..147  
 FT /note= "CL3"  
 FT Region 137..142  
 FT /note= "amphipathic helix"  
 FT Region 139..154  
 FT /note= "SA11"  
 FT Region 148..154  
 FT /note= "CL4"  
 XX  
 XX US5188936-A.  
 XX  
 PD 23-FEB-1993.  
 XX  
 PF 16-JAN-1991; 91US-00641346.  
 XX  
 PR 16-JAN-1991; 91US-00641346.  
 XX  
 PA (USDA ) US SEC OF AGRIC.  
 PA (IOWA ) UNIV IOWA STATE RES FOUND.  
 XX  
 PI Tabatabai LB, Mayfield JE, Beck BL;  
 XX  
 DR WPI; 1993-085536/10.  
 XX  
 FT Detection of Brucella abortus antibody - using B. abortus copper-zinc

PT superoxidedismutase protein or segments contg. antigenic determinants.  
 XX Disclosure; Fig 1; 12pp; English.  
 XX  
 CC The sequence is that of the 20kD Brucella abortus copper-zinc superoxide  
 CC dismutase (CuZnSOD) which is used as part of a method for detecting B.  
 CC abortus infection in animals, in particular Bovine brucellosis. The  
 CC method can distinguish between animals which have a natural infection and  
 CC those which have been vaccinated. CuZnSOD or a segment effective as an  
 CC antigenic determinant is combined with a body fluid sample and the  
 CC presence of a complex of the CuZnSOD and antibody detected. (Updated on  
 CC 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI  
 CC field.) (Updated on 24-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 154 AA;  
 Query Match 43.5%; Score 413.5; DB 2; Length 154;  
 Best Local Similarity 52.0%; Pred. No. 1.4e-38;  
 Matches 79; Conservative 19; Mismatches 53; Indels 1; Gaps 1;  
 QY 28 SVTSEVHMIDDNGIKQSIGTVTFTDTDKGLQIKTDKGLPAGHGFHIEGSGCPAEBHD 87  
 Db 2 STTVKMYEALPTGPGKEVGTTVVISEAPGGLHFKVNMKLTGPHGHVHENSAPCEKD 61  
 QY 88 GHLTAGLQAHGHYDPDKTKHEGPIGNGHKGDLPLRVKADGIAKETLLAPRL-TVKKEIK 146  
 Db 62 GKIVPALAAGGHYDPGNTHHHLGPEGDGHMGDLPLRLSANADGKVSFTVVAHLKLAIEIK 121  
 QY 147 GRTVMIHAGGDNYSKPLPLGGGGARIAICGVI 178  
 Db 122 QRLSMVHVGGDNYSKPEPLGGGGARFACGVI 153

RESULT 5  
 ADC00874  
 ID ADC00874 standard; protein; 175 AA.  
 XX  
 AC ADC00874;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Enterohaemorrhagic E. coli 0157:H7-specific protein SEQ ID NO: 919.  
 XX  
 KW enterohaemorrhagic; anti-bacterial.  
 XX  
 OS Escherichia coli; 0157:H7.  
 XX  
 PN JP2002355074-A.  
 XX  
 PD 10-DEC-2002.  
 XX  
 PF 24-JAN-2002; 2002JP-00015959.  
 XX  
 PR 24-JAN-2001; 2001JP-00112010.  
 XX  
 PA (UYTS-) UNIV TSUKUBA.  
 XX  
 XX WPI; 2003-451640/43.  
 XX  
 PT Enterohaemorrhagic Escherichia coli 0157:H7-specific nucleic acid molecule  
 PT and a polypeptide and its use, a polypeptide, a vector and a host cell.  
 XX  
 PS Claim 3; SEQ ID NO 919; 2067pp; Japanese.  
 XX  
 CC The invention relates to a novel enterohaemorrhagic Escherichia coli  
 CC 0157:H7-specific nucleic acid molecule. A polynucleotide of the invention  
 CC has anti-bacterial activity. The polypeptide can be used in detection  
 CC and/or treatment of 0157:H7 infection. The nucleotide sequence of the  
 CC genome of Enterohaemorrhagic E coli 0157:H7 was determined. The present  
 CC sequence represents an E. coli 0157:H7-specific polypeptide of the  
 CC invention.  
 XX  
 SQ Sequence 175 AA;



Db 68 HGFHIIHANGSCPEMDKMGKFPVPALKAGGHLDPENKGVHLGPNYKNGHGLDPLGLVANSKG 127  
QY 130 IAKETLLAPRLT-VKEIKGRTVMIHAGGDNYSKPLPLGGGGARIACGVI 178  
Db 128 DADYAVLAPRLTKLDQIKKALMVHVGGDNYSNPEALGGGGARMACGVI 177

RESULT 8  
ID ABB64175  
XX ABB64175 standard; protein; 179 AA.  
AC ABB64175;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 19317.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
WPI; 2001-656860/75.  
DR N-PSDB; ABL08278.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.  
XX  
PS Disclosure; SEQ ID NO 19317; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

QY 11 IVTISLTSITSVVLACSVTSEVHMI-----DDNGIKQSIGVTFTDTDKG--LQIKTDL 63  
Db 6 VWSLALCATICSAAQTMPICAIAYLIGPVQSDNTQVK---GNVFTQNDGQNVHVRVL 62  
QY 64 KGLPAGEHGFHIEGSCGPAEDHGLTAGLQAHG-HYDPDKTGKHEGLGN-GHKDLP 121  
Db 63 EGLKEKGHGFHIE-----KGLTNGCISMGAHYNPDKV-DRGGPDHEVHRVGDIG 112  
QY 122 RLVVKADGLAKETLLAPRLTVK---EIKGRTVMIH-----AGDNYSDKPLPLGGGGARI 173  
Db 113 NLEANSGLIDVITYDQVITLTKGLIGRGVVVHELEDLGLGNHTDSK-KTGNAGGRI 171

Query Match 21.8%; Score 207.5; DB 4; Length 179;  
Best Local Similarity 34.6%; Pred. No. 5.1e-15;  
Matches 64; Conservative 25; Mismatches 65; Indels 31; Gaps 10;

QY 16 LTTSTSVVLACSVTSEVHMI-----IDNGIKQSIGVTFTDTDKG--LQIK 60  
Db 5 LVVSLALCATICSAAQTRNMPICAIAYLIGPVQSDNTQVK---GNVFTQNDGQNVHVR 61  
QY 61 TDLKGLPAGEHGFHIEGSCGPAEDHGLTAGLQAHG-HYDPDKTGKHEGLGN-GHKG 118  
Db 62 VOLEGLKEKGHGFHIE-----KGLTNGCISMGAHYNPDKV-DRGGPDHEVHRVHG 111  
QY 119 DLPLRVKADGLAKETLLAPRLTVK---EIKGRTVMIH-----AGDNYSDKPLPLGGGG 170  
Db 112 DLGNLEANSGLIDVITYDQVITLTKGLIGRGVVVHELEDLGLGNHTDSK-KTGNAG 170  
QY 171 ARIACGVI 178  
|||||

QY 174 ACGVI 178  
|||||  
Db 172 ACGVI 176  
|||||

RESULT 9  
ID ABB67296  
XX ABB67296 standard; protein; 181 AA.  
AC ABB67296;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 28680.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
WPI; 2001-656860/75.  
DR N-PSDB; ABL1399.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.  
XX  
PS Disclosure; SEQ ID NO 28680; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

QY 16 LTTSTSVVLACSVTSEVHMI-----IDNGIKQSIGVTFTDTDKG--LQIK 60  
Db 5 LVVSLALCATICSAAQTRNMPICAIAYLIGPVQSDNTQVK---GNVFTQNDGQNVHVR 61  
QY 61 TDLKGLPAGEHGFHIEGSCGPAEDHGLTAGLQAHG-HYDPDKTGKHEGLGN-GHKG 118  
Db 62 VOLEGLKEKGHGFHIE-----KGLTNGCISMGAHYNPDKV-DRGGPDHEVHRVHG 111  
QY 119 DLPLRVKADGLAKETLLAPRLTVK---EIKGRTVMIH-----AGDNYSDKPLPLGGGG 170  
Db 112 DLGNLEANSGLIDVITYDQVITLTKGLIGRGVVVHELEDLGLGNHTDSK-KTGNAG 170  
QY 171 ARIACGVI 178  
|||||

Query Match 20.7%; Score 196.5; DB 4; Length 181;  
Best Local Similarity 34.0%; Pred. No. 9.4e-14;  
Matches 64; Conservative 19; Mismatches 66; Indels 39; Gaps 10;





PS Disclosure; Col 77-88; 140pp; English.

XX The present invention relates to selenoproteins encoded in the genome of a virus, where the coding sequence of the selenoprotein is genetically engineered for expression in a nucleic acid construct. The invention also discloses a method for identifying selenoprotein coding sequences, for detecting certain viruses (e.g. HIV or Ebola), cancer and immune system disorders. The present sequence was used to illustrate the invention

XX Sequence 202 AA;

Query Match 18.7%; Score 178; DB 5; Length 202;  
Best Local Similarity 29.3%; Pred. No. 1.4e-11;  
Matches 53; Conservative 23; Mismatches 59; Indels 46; Gaps 9;

QY 17 LTSITSVVLACVTSEVHMIDNGIKQSIGTFTDTDKG-LQIKTDLKGLPAGBHGHI 75  
DB 44 LTVAAAKKAVSVLKGTSV-----GVTLTQDDDEGPTTVNRIITGLTFGLHGFHL 95  
QY 76 HEGG-----SCGPAEHDGHLTAGLQAHGHYDPDKTGRKEGFLGN-GHKGDLPRLVVKRA 127  
DB 96 HEYGDITNGCISTGP-----HFNPNKL-THGAPDEIRHAGDLGNIVANA 139  
QY 128 DGIAXETLL---APRLTVKEIKGRVTMIIH-----AGDNYSDKPLPLGGGARIACGV 177  
DB 140 EGVABATIVDNGIPLTGTNSVVGRLVVLHQLDGLKGGHELS---LSTGNAGGRLACGV 196  
QY 178 I 178  
DB 197 V 197

RESULT 15  
AAM52487  
ID AAM52487 standard; protein; 150 AA.  
XX AAM52487;  
AC AAM52487;  
XX  
DT 03-JUL-2002 (first entry)  
DE Superoxide dismutase protein #16.  
XX Selenoprotein; HIV; Ebola virus; cancer; immune system disorder.  
XX Xenopus laevis.  
OS  
PN US6303295-B1.  
XX  
PD 16-OCT-2001.  
XX  
PF 12-JUL-1996; 96US-00679493.  
XX  
PR 14-JUL-1995; 95US-0001203P.  
PR 01-SEP-1995; 95US-0003112P.  
XX  
PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX  
PI Taylor EW, Nadingpalli RG, Ramanathan CS;  
XX  
PN US6303295-B1.  
XX  
PD 16-OCT-2001.  
XX  
PF 12-JUL-1996; 96US-00679493.  
XX  
PR 14-JUL-1995; 95US-0001203P.  
PR 01-SEP-1995; 95US-0003112P.  
XX  
PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX  
PI Taylor EW, Nadingpalli RG, Ramanathan CS;  
XX  
PN WPI; 2002-024734/03.  
XX  
PT New selenoprotein for use in detecting certain viruses, e.g. human immunodeficiency virus (HIV) or Ebola, cancer and immune system disorders.  
PS Disclosure; Col 77-88; 140pp; English.

XX The present invention relates to selenoproteins encoded in the genome of a virus, where the coding sequence of the selenoprotein is genetically engineered for expression in a nucleic acid construct. The invention also discloses a method for identifying selenoprotein coding sequences, for detecting certain viruses (e.g. HIV or Ebola), cancer and immune system disorders. The present sequence was used to illustrate the invention

XX SQ Sequence 150 AA;

Query Match 18.5%; Score 175.5; DB 5; Length 150;  
Best Local Similarity 33.8%; Pred. No. 1.8e-11;  
Matches 50; Conservative 18; Mismatches 49; Indels 31; Gaps 7;

QY 46 GTVTFTDTDKG-LQIKTDLKGLPAGBHGHIHEGSCGPAEHDGHLTAG-----LQAHGHY 100  
DB 15 GVVREFEQDDGDVTVEGKIEGLTDGNHGFHIVFGD-----NTNGCLSAGPHFNQKNKG 70  
QY 101 DPKDTGRKEGFLGNHGKGDLPRLVVKADGIAKETILLAPRLTVK---EIKGRVTMIIH--- 153  
DB 71 SPKADRHVGDIGN-----VTAEGGVAQFKTDPQISLKGRSIIIGRTAVVHEKQD 121  
QY 154 ---AGDNYSDKPLPLGGGARIACGV 178  
DB 122 DLGRGGD---DESLKTGNAGGRLACGV 146

RESULT 16  
AAM52497  
ID AAM52497 standard; protein; 166 AA.  
XX AAM52497;  
AC AAM52497;  
XX  
DT 03-JUL-2002 (first entry)  
DE Superoxide dismutase protein #26.  
XX Selenoprotein; HIV; Ebola virus; cancer; immune system disorder.  
XX Unidentified.  
OS  
PN US6303295-B1.  
XX  
PD 16-OCT-2001.  
XX  
PF 12-JUL-1996; 96US-00679493.  
XX  
PR 14-JUL-1995; 95US-0001203P.  
PR 01-SEP-1995; 95US-0003112P.  
XX  
PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX  
PI Taylor EW, Nadingpalli RG, Ramanathan CS;  
XX  
PN WPI; 2002-024734/03.  
XX  
PT New selenoprotein for use in detecting certain viruses, e.g. human immunodeficiency virus (HIV) or Ebola, cancer and immune system disorders.  
PS Disclosure; Col 77-88; 140pp; English.

XX The present invention relates to selenoproteins encoded in the genome of a virus, where the coding sequence of the selenoprotein is genetically engineered for expression in a nucleic acid construct. The invention also discloses a method for identifying selenoprotein coding sequences, for detecting certain viruses (e.g. HIV or Ebola), cancer and immune system disorders. The present sequence was used to illustrate the invention

XX SQ Sequence 166 AA;

Query Match 18.5%; Score 175.5; DB 5; Length 166;  
Best Local Similarity 33.1%; Pred. No. 2.1e-11;  
Matches 50; Conservative 31; Mismatches 45; Indels 25; Gaps 11;

QY 41 IKQSI-GRVTFDTDKG-LQIKTDLKGLPAGBHGHIHEGSCGPAEHDGHLTAGLQAHG 98  
DB 23 VKPEVKGIIEYEQQNGVTLTSGSITGLTEKGHGFHVHEFGD-----NTNGCTSAG---A 74  
QY 99 HYDPDKTGR-HEGFLGN-GHKGDLPRLVVKADGIA-----KETLLAPRLT-VKEIKGRVTM 151







|    |              |                |    |              |                |
|----|--------------|----------------|----|--------------|----------------|
| PR | 07-JUN-1999; | 99US-0137724P. | PR | 06-AUG-1999; | 99US-0147416P. |
| PR | 08-JUN-1999; | 99US-0138094P. | PR | 09-AUG-1999; | 99US-0147433P. |
| PR | 10-JUN-1999; | 99US-0138540P. | PR | 09-AUG-1999; | 99US-0147935P. |
| PR | 10-JUN-1999; | 99US-0138847P. | PR | 10-AUG-1999; | 99US-0148171P. |
| PR | 14-JUN-1999; | 99US-0139119P. | PR | 11-AUG-1999; | 99US-0148319P. |
| PR | 16-JUN-1999; | 99US-0139452P. | PR | 12-AUG-1999; | 99US-0148341P. |
| PR | 16-JUN-1999; | 99US-0139453P. | PR | 13-AUG-1999; | 99US-0148565P. |
| PR | 17-JUN-1999; | 99US-0139492P. | PR | 13-AUG-1999; | 99US-0148684P. |
| PR | 18-JUN-1999; | 99US-0139454P. | PR | 16-AUG-1999; | 99US-0149368P. |
| PR | 18-JUN-1999; | 99US-0139455P. | PR | 17-AUG-1999; | 99US-0149175P. |
| PR | 18-JUN-1999; | 99US-0139456P. | PR | 18-AUG-1999; | 99US-0149426P. |
| PR | 18-JUN-1999; | 99US-0139457P. | PR | 20-AUG-1999; | 99US-0149722P. |
| PR | 18-JUN-1999; | 99US-0139458P. | PR | 20-AUG-1999; | 99US-0149723P. |
| PR | 18-JUN-1999; | 99US-0139459P. | PR | 20-AUG-1999; | 99US-0149929P. |
| PR | 18-JUN-1999; | 99US-0139460P. | PR | 23-AUG-1999; | 99US-0149902P. |
| PR | 18-JUN-1999; | 99US-0139461P. | PR | 23-AUG-1999; | 99US-0149930P. |
| PR | 18-JUN-1999; | 99US-0139462P. | PR | 25-AUG-1999; | 99US-0150566P. |
| PR | 18-JUN-1999; | 99US-0139463P. | PR | 26-AUG-1999; | 99US-0150884P. |
| PR | 18-JUN-1999; | 99US-0139750P. | PR | 27-AUG-1999; | 99US-0151065P. |
| PR | 18-JUN-1999; | 99US-0139763P. | PR | 27-AUG-1999; | 99US-0151066P. |
| PR | 21-JUN-1999; | 99US-0139817P. | PR | 27-AUG-1999; | 99US-0151080P. |
| PR | 22-JUN-1999; | 99US-0139899P. | PR | 30-AUG-1999; | 99US-0151303P. |
| PR | 23-JUN-1999; | 99US-0140353P. | PR | 31-AUG-1999; | 99US-0151438P. |
| PR | 23-JUN-1999; | 99US-0140354P. | PR | 01-SEP-1999; | 99US-0151930P. |
| PR | 24-JUN-1999; | 99US-0140695P. | PR | 07-SEP-1999; | 99US-0152363P. |
| PR | 28-JUN-1999; | 99US-0140823P. | PR | 10-SEP-1999; | 99US-0153070P. |
| PR | 29-JUN-1999; | 99US-0140991P. | PR | 13-SEP-1999; | 99US-0153758P. |
| PR | 30-JUN-1999; | 99US-0141287P. | PR | 15-SEP-1999; | 99US-0154018P. |
| PR | 01-JUL-1999; | 99US-0141842P. | PR | 16-SEP-1999; | 99US-0154039P. |
| PR | 01-JUL-1999; | 99US-0142154P. | PR | 20-SEP-1999; | 99US-0154779P. |
| PR | 02-JUL-1999; | 99US-0142055P. | PR | 22-SEP-1999; | 99US-0155139P. |
| PR | 06-JUL-1999; | 99US-0142390P. | PR | 23-SEP-1999; | 99US-0155486P. |
| PR | 08-JUL-1999; | 99US-0142803P. | PR | 24-SEP-1999; | 99US-0155659P. |
| PR | 09-JUL-1999; | 99US-0142920P. | PR | 28-SEP-1999; | 99US-0156458P. |
| PR | 12-JUL-1999; | 99US-0142977P. | PR | 29-SEP-1999; | 99US-0156596P. |
| PR | 13-JUL-1999; | 99US-0143342P. | PR | 04-OCT-1999; | 99US-0157117P. |
| PR | 14-JUL-1999; | 99US-0143624P. | PR | 05-OCT-1999; | 99US-0157753P. |
| PR | 15-JUL-1999; | 99US-0144005P. | PR | 06-OCT-1999; | 99US-0157865P. |
| PR | 16-JUL-1999; | 99US-0144085P. | PR | 07-OCT-1999; | 99US-0158029P. |
| PR | 16-JUL-1999; | 99US-0144086P. | PR | 08-OCT-1999; | 99US-0158232P. |
| PR | 19-JUL-1999; | 99US-0144325P. | PR | 12-OCT-1999; | 99US-0158369P. |
| PR | 19-JUL-1999; | 99US-0144331P. | PR | 13-OCT-1999; | 99US-0159294P. |
| PR | 19-JUL-1999; | 99US-0144332P. | PR | 13-OCT-1999; | 99US-0159295P. |
| PR | 19-JUL-1999; | 99US-0144333P. | PR | 14-OCT-1999; | 99US-0159329P. |
| PR | 19-JUL-1999; | 99US-0144334P. | PR | 14-OCT-1999; | 99US-0159330P. |
| PR | 19-JUL-1999; | 99US-0144335P. | PR | 14-OCT-1999; | 99US-0159331P. |
| PR | 20-JUL-1999; | 99US-0144352P. | PR | 14-OCT-1999; | 99US-0159637P. |
| PR | 20-JUL-1999; | 99US-0144632P. | PR | 14-OCT-1999; | 99US-0159638P. |
| PR | 21-JUL-1999; | 99US-0144884P. | PR | 18-OCT-1999; | 99US-0159584P. |
| PR | 21-JUL-1999; | 99US-0144814P. | PR | 21-OCT-1999; | 99US-0160741P. |
| PR | 21-JUL-1999; | 99US-0145086P. | PR | 21-OCT-1999; | 99US-0160767P. |
| PR | 21-JUL-1999; | 99US-0145088P. | PR | 21-OCT-1999; | 99US-0160768P. |
| PR | 22-JUL-1999; | 99US-0145085P. | PR | 21-OCT-1999; | 99US-0160770P. |
| PR | 22-JUL-1999; | 99US-0145087P. | PR | 21-OCT-1999; | 99US-0160814P. |
| PR | 22-JUL-1999; | 99US-0145089P. | PR | 21-OCT-1999; | 99US-0160815P. |
| PR | 22-JUL-1999; | 99US-0145192P. | PR | 22-OCT-1999; | 99US-0160980P. |
| PR | 23-JUL-1999; | 99US-0145145P. | PR | 22-OCT-1999; | 99US-0160981P. |
| PR | 23-JUL-1999; | 99US-0145218P. | PR | 22-OCT-1999; | 99US-0160981P. |
| PR | 23-JUL-1999; | 99US-0145224P. | PR | 22-OCT-1999; | 99US-0160989P. |
| PR | 26-JUL-1999; | 99US-0145276P. | PR | 25-OCT-1999; | 99US-0161404P. |
| PR | 27-JUL-1999; | 99US-0145913P. | PR | 25-OCT-1999; | 99US-0161405P. |
| PR | 27-JUL-1999; | 99US-0145918P. | PR | 25-OCT-1999; | 99US-0161406P. |
| PR | 27-JUL-1999; | 99US-0145919P. | PR | 26-OCT-1999; | 99US-0161359P. |
| PR | 28-JUL-1999; | 99US-0145951P. | PR | 26-OCT-1999; | 99US-0161360P. |
| PR | 02-AUG-1999; | 99US-0146386P. | PR | 26-OCT-1999; | 99US-0161361P. |
| PR | 02-AUG-1999; | 99US-0146388P. | PR | 28-OCT-1999; | 99US-0161920P. |
| PR | 02-AUG-1999; | 99US-0146389P. | PR | 28-OCT-1999; | 99US-0161992P. |
| PR | 03-AUG-1999; | 99US-0147038P. | PR | 28-OCT-1999; | 99US-0161993P. |
| PR | 04-AUG-1999; | 99US-0147204P. | PR | 29-OCT-1999; | 99US-0162142P. |
| PR | 04-AUG-1999; | 99US-0147302P. |    |              |                |
| PR | 05-AUG-1999; | 99US-0147192P. |    |              |                |
| PR | 05-AUG-1999; | 99US-0147260P. |    |              |                |
| PR | 06-AUG-1999; | 99US-0147303P. |    |              |                |

Query Match 18.4%; Score 175; DB 3; Length 216;  
Best Local Similarity 30.1%; Pred. No. 3.4e-11;  
Matches 55; Conservative 22; Mismatches 56; Indels 50; Gaps 9;



PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 21-MAY-1999; 99US-0135529P.  
PR 24-MAY-1999; 99US-0136021P.  
PR 25-MAY-1999; 99US-0136392P.  
PR 27-MAY-1999; 99US-0136782P.  
PR 28-MAY-1999; 99US-0137222P.  
PR 01-JUN-1999; 99US-0137528P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 10-JUN-1999; 99US-0139119P.  
PR 14-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0145868P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149029P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.

|  |  |  |    |              |                |
|--|--|--|----|--------------|----------------|
| PR   | 25-OCT-1999;   | 99US-0161404P.   | PR | 06-MAY-1999; | 99US-0132486P. |
| PR   | 25-OCT-1999;   | 99US-0161405P.   | PR | 06-MAY-1999; | 99US-0132487P. |
| PR   | 25-OCT-1999;   | 99US-0161406P.   | PR | 07-MAY-1999; | 99US-0132863P. |
| PR   | 26-OCT-1999;   | 99US-0161359P.   | PR | 11-MAY-1999; | 99US-0134258P. |
| PR   | 26-OCT-1999;   | 99US-0161360P.   | PR | 14-MAY-1999; | 99US-0134218P. |
| PR   | 26-OCT-1999;   | 99US-0161361P.   | PR | 14-MAY-1999; | 99US-0134219P. |
| PR   | 28-OCT-1999;   | 99US-0161920P.   | PR | 14-MAY-1999; | 99US-0134221P. |
| PR   | 28-OCT-1999;   | 99US-0161992P.   | PR | 14-MAY-1999; | 99US-0134370P. |
| PR   | 28-OCT-1999;   | 99US-0161993P.   | PR | 18-MAY-1999; | 99US-0134769P. |
| PR   | 29-OCT-1999;   | 99US-0162142P.   | PR | 19-MAY-1999; | 99US-0134941P. |
| Query Match 18.4%; Score 175; DB 3; Length 230;                |  |  |    |              |                |
| Best Local Similarity 30.1%; Pred. No. 3.7e-11;                |  |  |    |              |                |
| Matches 55; Conservative 22; Mismatches 56; Indels 50; Gaps 9; |  |  |    |              |                |
| QY   | 17   | LTSITSVVLACSV--TSEVHMIDDNGIKQSGTGTFTDQK-LQIKTDLKLPGAHG 72        | PR | 01-JUN-1999; | 99US-0137222P. |
| Db   | 72   | LTVSAAKXAVALKGTSDE-----GVTLTQDDSGPTTVNRITGLTFPGPH 120            | PR | 03-JUN-1999; | 99US-0137528P. |
| QY   | 73   | FHIHEGG-----SCGPAEHDGHLTAGLQAHGCHYDPDKTKHKGFLNGHKGDLFLVLV 125    | PR | 04-JUN-1999; | 99US-0137502P. |
| Db   | 121  | PHLHEFGDTNGCISGPHFNENMT-----HGAPEDECRHAGDLGN-----INA 165         | PR | 07-JUN-1999; | 99US-0137724P. |
| QY   | 126  | KADGIAXETLL---APRLTVKEIKGRTVMIH-----AGGNYSDKPLPLGGGGARIAC 175    | PR | 08-JUN-1999; | 99US-0138094P. |
| Db   | 166  | NADGVAETTTVDNQIFLTGPNVSVGAFVYVHELKDDLKGGGHELS---LTTGNAGGRLAC 222 | PR | 10-JUN-1999; | 99US-0138540P. |
| QY   | 176  | GVI 178  | PR | 10-JUN-1999; | 99US-0138847P. |
| Db   | 223  | GVI 225  | PR | 14-JUN-1999; | 99US-0139119P. |
| RESULT 21  |  |  |    |              |                |
| AAG05963   |  |  |    |              |                |
| ID   | AAG05963 standard; protein; 232 AA.                                      |  |    |              |                |
| XX   | AC AAG05963;   |  |    |              |                |
| XX   | 17-OCT-2000 (first entry)  |  |    |              |                |
| DE   | Arabidopsis thaliana protein fragment SEQ ID NO: 2562.                   |  |    |              |                |
| KW   | Protein identification; signal transduction pathway; metabolic pathway;  |  |    |              |                |
| KW   | hybridisation assay; genetic mapping; gene expression control; promoter; |  |    |              |                |
| KW   | termination sequence.  |  |    |              |                |
| OS   | Arabidopsis thaliana.  |  |    |              |                |
| PN   | EP1033405-A2.  |  |    |              |                |
| XX   | 06-SEP-2000.   |  |    |              |                |
| XX   | 25-FEB-2000; 2000EP-00301439.  |  |    |              |                |
| PR   | 25-FEB-1999;   | 99US-0121823P.   | PR | 08-JUL-1999; | 99US-0142803P. |
| PR   | 05-MAR-1999;   | 99US-0123180P.   | PR | 09-JUL-1999; | 99US-0142920P. |
| PR   | 09-MAR-1999;   | 99US-0123548P.   | PR | 12-JUL-1999; | 99US-0142977P. |
| PR   | 23-MAR-1999;   | 99US-0125789P.   | PR | 13-JUL-1999; | 99US-0143542P. |
| PR   | 25-MAR-1999;   | 99US-0126264P.   | PR | 14-JUL-1999; | 99US-0143624P. |
| PR   | 29-MAR-1999;   | 99US-0126785P.   | PR | 15-JUL-1999; | 99US-0144005P. |
| PR   | 01-APR-1999;   | 99US-0127462P.   | PR | 16-JUL-1999; | 99US-0144085P. |
| PR   | 06-APR-1999;   | 99US-0128234P.   | PR | 19-JUL-1999; | 99US-0144325P. |
| PR   | 08-APR-1999;   | 99US-0128714P.   | PR | 19-JUL-1999; | 99US-0144332P. |
| PR   | 16-APR-1999;   | 99US-0129845P.   | PR | 19-JUL-1999; | 99US-0144333P. |
| PR   | 19-APR-1999;   | 99US-0130077P.   | PR | 19-JUL-1999; | 99US-0144334P. |
| PR   | 21-APR-1999;   | 99US-0130449P.   | PR | 20-JUL-1999; | 99US-0144335P. |
| PR   | 23-APR-1999;   | 99US-0130510P.   | PR | 20-JUL-1999; | 99US-0144632P. |
| PR   | 23-APR-1999;   | 99US-0130891P.   | PR | 20-JUL-1999; | 99US-0144633P. |
| PR   | 28-APR-1999;   | 99US-0131449P.   | PR | 21-JUL-1999; | 99US-0144814P. |
| PR   | 30-APR-1999;   | 99US-0132048P.   | PR | 21-JUL-1999; | 99US-0145086P. |
| PR   | 30-APR-1999;   | 99US-0132407P.   | PR | 21-JUL-1999; | 99US-0145088P. |
| PR   | 04-MAY-1999;   | 99US-0132484P.   | PR | 22-JUL-1999; | 99US-0145085P. |
| PR   | 05-MAY-1999;   | 99US-0132485P.   | PR | 22-JUL-1999; | 99US-0145087P. |



```

XX Tainer JA, Kuhn L, Boissinot M, Fisher C, Farge HE, Griffin JH;
PI Mullenbach GT, Hallelwell RA;
XX WPI: 1992-183671/22.
DR P-PSDB; AAR24226.
XX
XX Fusion proteins with glycosamino:glycan-binding and superoxidisedismutase
PT activities - reduce tissue damage caused by super:oxide radicals, useful
PT in treating autoimmune diseases e.g. rheumatoid arthritis and
PT osteoarthritis.
XX
XX Example 1; Fig 1; 140pp; English.
XX
XX The fusion protein was constructed to contain the heparin binding region
CC of Protein C inhibitor (PCI), namely the A+ amphipathic alpha helix of
CC PCI, and human superoxide dismutase, joined via a linker region. The
CC fusion protein is useful for extending the in vivo lifetimes of
CC biologically active cpds. such as SOD and for targeting them to specific
CC cell surfaces or substrates. The glycosaminoglycan (GAG) binding protein
CC is formed into a fusion protein with SOD to increase stability, plasma
CC half-life and ease of purification of SOD. SOD is useful for reduction of
CC tissue damage caused by oxygen radicals and is used in the treatment of
CC autoimmune diseases e.g. rheumatoid and osteo-arthritis. See also
CC AAR24226-35, AAR27932-51. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 196 AA;
SQ
Query Match 18.2%; Score 172.5; DB 2; Length 196;
Best Local Similarity 31.7%; Pred. No. 5.7e-11;
Matches 53; Conservative 26; Mismatches 59; Indels 29; Gaps 9;
QY 26 ACSVTSEVHMIDNGIKQSIGTVTFTDTKGLQIKT--DLKGLPAGEHGFIHEGSGCP 83
DB 20 ALAATKAVAVLKGDGPVQ--GIINFEQKESNGPVKVGSIKGLTGLGHGFHVEFGD--- 74
QY 84 AEHDGHLTAGLQAHG-HYDPPDKTKHGEPLG-NGHKGDPLRLVVKADGIKAKETLLAPRLT 141
DB 75 -----NTAGCTSAGPHFNP-LSRXHGGPKDEERHVGDLGNVTADKGVADVSIEDSVIS 127
QY 142 VK---EIKGRTVMIH-----AGDNYSDKPLPLGGGGARIACGVI 178
DB 128 LSGDHSIIIGRTLIVVHEKADDLGKGNEESTK---TGNAGSRKACGVI 171
RESULT 23
AAR27948
ID AAR27948 standard; protein; 183 AA.
XX
XX AAR27948;
AC
XX
XX 25-MAR-2003 (revised)
DT 25-NOV-1992 (first entry)
XX
XX GAG fusion protein with SOD according to a formula.
DE
XX Glycosamino:glycan; superoxidisedismutase; tissue damage;
KW autoimmune disease; rheumatoid arthritis; osteoarthritis; ss.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Peptide 1..11
FT /note= "GAG binding motif"
FT Peptide 12..183
FT /note= "SOD "
XX
XX WO9207935-A1.
PN
XX 14-MAY-1992.
PD
XX
XX 01-NOV-1991; 91WO-US008105.
PF

```

```

XX 01-NOV-1990; 90US-00608539.
PR 02-NOV-1990; 90US-00608569.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Tainer JA, Kuhn L, Boissinot M, Fisher C, Farge HE, Griffin JH;
PI Mullenbach GT, Hallelwell RA;
XX
XX WPI: 1992-183671/22.
DR
XX
XX Fusion proteins with glycosamino:glycan-binding and superoxidisedismutase
PT activities - reduce tissue damage caused by super:oxide radicals, useful
PT in treating autoimmune diseases e.g. rheumatoid arthritis and
PT osteoarthritis.
XX
XX Claim 8; Fig 1; 140pp; English.
XX
XX The fusion protein comprising the a glycosaminoglycan binding region and
CC human superoxide dismutase, joined via a linker region was constructed
CC according to the formula M-(Z-M)-SOD where Z is the peptide -EKLTKWLK-.
CC The fusion protein is useful for extending the in vivo lifetimes of
CC biologically active cpds. such as SOD and for targeting them to specific
CC cell surfaces or substrates. The glycosaminoglycan (GAG) binding protein
CC is formed into a fusion protein with SOD to increase stability, plasma
CC half-life and ease of purification of SOD. SOD is useful for reduction of
CC tissue damage caused by oxygen radicals and is used in the treatment of
CC autoimmune diseases e.g. rheumatoid and osteo-arthritis. See also
CC AAR24225-35, AAR27932-51. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 183 AA;
SQ
Query Match 18.1%; Score 171.5; DB 2; Length 183;
Best Local Similarity 31.5%; Pred. No. 6.7e-11;
Matches 53; Conservative 26; Mismatches 60; Indels 29; Gaps 9;
QY 25 LACSVTSEVHMIDNGIKQSIGTVTFTDTKGLQIKT--DLKGLPAGEHGFIHEGSGCG 82
DB 9 LKMAATKAVAVLKGDGPVQ--GIINFEQKESNGPVKVGSIKGLTGLGHGFHVEFGD--- 64
QY 83 PAHDGHLTAGLQAHG-HYDPPDKTKHGEPLG-NGHKGDPLRLVVKADGIKAKETLLAPRL 140
DB 65 -----NTAGCTSAGPHFNP-LSRXHGGPKDEERHVGDLGNVTADKGVADVSIEDSVI 116
QY 141 TVK---EIKGRTVMIH-----AGDNYSDKPLPLGGGGARIACGVI 178
DB 117 LSGDHSIIIGRTLIVVHEKADDLGKGNEESTK---TGNAGSRKACGVI 161
RESULT 24
AAR27938
ID AAR27938 standard; protein; 185 AA.
XX
XX AAR27938;
AC
XX
XX 25-MAR-2003 (revised)
DT 25-NOV-1992 (first entry)
XX
XX GAG fusion protein with SOD according to a formula.
DE
XX Glycosamino:glycan; superoxidisedismutase; tissue damage;
KW autoimmune disease; rheumatoid arthritis; osteoarthritis; ss.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Peptide 1..13
FT /note= "GAG binding motif"
FT Peptide 14..185
FT /note= "SOD "
XX
XX WO9207935-A1.
PF

```



```
XX 14-MAY-1992.
XX
XX 01-NOV-1991; 91WO-US008105.
XX
XX 01-NOV-1990; 90US-00608539.
XX
XX 02-NOV-1990; 90US-00608569.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Tainer JA, Kuhn L, Boissinot M, Fisher C, Parge HE, Griffin JH;
XX Mullenbach GT, Hallewell RA;
XX WPI; 1992-183671/22.
XX
XX Fusion proteins with glycosamino:glycan-binding and superoxidisedismutase
XX activities - reduce tissue damage caused by super:oxide radicals, useful
XX in treating autoimmune diseases e.g. rheumatoid arthritis and
XX osteoarthritis.
XX
XX Claim 8; Fig 1; 140pp; English.
XX
XX The fusion protein comprising the a glycosaminoglycan binding region and
XX human superoxide dismutase, joined via a linker region was constructed
XX according to the formula M-(Z-M)-SOD where Z is the peptide -YKKIKKLES-
XX . The fusion protein is useful for extending the in vivo lifetimes of
XX biologically active cpds. such as SOD and for targeting them to specific
XX cell surfaces or substrates. The glycosaminoglycan (GAG) binding protein
XX is formed into a fusion protein with SOD to increase stability, plasma
XX half-life and ease of purification of SOD. SOD is useful for reduction of
XX tissue damage caused by oxygen radicals and is used in the treatment of
XX autoimmune diseases e.g. rheumatoid and osteo-arthritis. See also
XX AAR24225-35, AAR27932-51. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 185 AA;
Query Match 18.1%; Score 171.5; DB 2; Length 185;
Best Local Similarity 30.6%; Pred. No. 6.8e-11;
Matches 53; Conservative 29; Mismatches 62; Indels 29; Gaps 9;
QY 20 ITSVLACSVTSEVHMIDDNGIKQSIGTFTDTDKGLQIKT--DLKGLPAGEHGFHIHE 77
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6 IKKLESMAATKAVAVLKGDPVQ--GIINFEQKESNGPVKVGSIKGLTEGLHGFHVHE 63
QY 78 GSGCGPAHGHCHLTAGLQAHG-HYDPDKTKHKGPLG-NGHKGDLPLRLVVKADGIKTEL 135
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
64 FGD-----NTAGCTSAGPHFNP-LSRKHGKPKDEERHVGDLGNVTADKDGVDVSI 113
QY 136 LAPRLTVK---EIKGRVTMTH-----AGDNYSDKPLPLGGGGARIACGVI 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
114 EDSVISLGDHSIIIGRTLNVHVKADDLKGGNEESTK---TGNAGSLACGVI 163
QY
Db
RESULT 25
AAR27947
ID AAR27947 standard; protein; 193 AA.
XX
XX AAR27947;
XX
XX 25-MAR-2003 (revised)
XX 25-NOV-1992 (first entry)
XX
XX GAG fusion protein with SOD according to a formula.
XX
XX Glycosamino:glycan; superoxidisedismutase; tissue damage;
XX autoimmune disease; rheumatoid arthritis; osteoarthritis; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 1..21
XX /note= "GAG binding motif"
XX
XX FT
```

```
FT Peptide 22..193
XX /note= "SOD "
XX
XX PN W09207935-A1.
XX
XX 14-MAY-1992.
XX
XX 01-NOV-1991; 91WO-US008105.
XX
XX 01-NOV-1990; 90US-00608539.
XX
XX 02-NOV-1990; 90US-00608569.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Tainer JA, Kuhn L, Boissinot M, Fisher C, Parge HE, Griffin JH;
XX Mullenbach GT, Hallewell RA;
XX WPI; 1992-183671/22.
XX
XX Fusion proteins with glycosamino:glycan-binding and superoxidisedismutase
XX activities - reduce tissue damage caused by super:oxide radicals, useful
XX in treating autoimmune diseases e.g. rheumatoid arthritis and
XX osteoarthritis.
XX
XX Claim 8; Fig 1; 140pp; English.
XX
XX The fusion protein comprising the a glycosaminoglycan binding region and
XX human superoxide dismutase, joined via a linker region was constructed
XX according to the formula M-(Z-M)-SOD where Z is the peptide -EKLRLKWLK-.
XX The fusion protein is useful for extending the in vivo lifetimes of
XX biologically active cpds. such as SOD and for targeting them to specific
XX cell surfaces or substrates. The glycosaminoglycan (GAG) binding protein
XX is formed into a fusion protein with SOD to increase stability, plasma
XX half-life and ease of purification of SOD. SOD is useful for reduction of
XX tissue damage caused by oxygen radicals and is used in the treatment of
XX autoimmune diseases e.g. rheumatoid and osteo-arthritis. See also
XX AAR24225-35, AAR27932-51. (Updated on 25-MAR-2003 to correct PN field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 193 AA;
Query Match 18.1%; Score 171.5; DB 2; Length 193;
Best Local Similarity 31.5%; Pred. No. 7.2e-11;
Matches 53; Conservative 26; Mismatches 60; Indels 29; Gaps 9;
QY 25 LACSVTSEVHMIDDNGIKQSIGTFTDTDKGLQIKT--DLKGLPAGEHGFHIHEGSCG 82
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
19 LKWAATKAVAVLKGDPVQ--GIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGD-- 74
QY 83 PAEHGDELTAGLQAHG-HYDPDKTKHKGPLG-NGHKGDLPLRLVVKADGIKTELAPRL 140
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
75 -----NTAGCTSAGPHFNP-LSRKHGKPKDEERHVGDLGNVTADKDGVDVSI 126
QY 141 TVK---EIKGRVTMTH-----AGDNYSDKPLPLGGGGARIACGVI 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 SLSGDHSIIIGRTLNVHVKADDLKGGNEESTK---TGNAGSLACGVI 171
QY
Db
RESULT 26
AAR27937
ID AAR27937 standard; protein; 197 AA.
XX
XX AAR27937;
XX
XX 25-MAR-2003 (revised)
XX 25-NOV-1992 (first entry)
XX
XX GAG fusion protein with SOD according to a formula.
XX
XX Glycosamino:glycan; superoxidisedismutase; tissue damage;
XX autoimmune disease; rheumatoid arthritis; osteoarthritis; ss.
XX
XX Synthetic.
XX
XX OS
```

```

XX Key Location/Qualifiers
FH Peptide 1..25
FT /note= "GAG binding motif"
FT Peptide 26..197
FT /note= "SOD"
XX
PN WO9207935-A1.
XX
PD 14-MAY-1992.
XX
PF 01-NOV-1991; 91WO-US008105.
XX
PR 01-NOV-1990; 90US-00608539.
PR 02-NOV-1990; 90US-00608569.
XX
PA (SCRI ) SCRIPPS RES INST.
XX
PI Tainer JA, Kuhn L, Boissinot M, Fisher C, Parge HE, Griffin JH;
PI Mullenbach GT, Hallewell RA;
XX
DR WPI; 1992-183671/22.
XX
PS Fusion proteins with glycosamino:glycan-binding and superoxidisedismutase
PS activities - reduce tissue damage caused by super:oxide radicals, useful
PS in treating autoimmune diseases e.g. rheumatoid arthritis and
PS osteoarthritis.
XX
PS Claim 8; Fig 1; 140pp; English.
XX
CC The fusion protein comprising the a glycosaminoglycan binding region and
CC human superoxide dismutase, joined via a linker region was constructed
CC according to the formula M-(Z-M)2-SOD where Z is the peptide -YKKIKKILLES
CC -. The fusion protein is useful for extending the in vivo lifetimes of
CC biologically active cpds. such as SOD and for targetting them to specific
CC cell surfaces or substrates. The glycosaminoglycan (GAG) binding protein
CC is formed into a fusion protein with SOD to increase stability, plasma
CC half-life and ease of purification of SOD. SOD is useful for reduction of
CC tissue damage caused by oxygen radicals and is used in the treatment of
CC autoimmune diseases e.g. rheumatoid and osteo-arthritis. See also
CC AAR24225-35, AAR27932-51. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 197 AA;
Query Match 18.1%; Score 171.5; DB 2; Length 197;
Best Local Similarity 30.6%; Pred. No. 7, 4e-11;
Matches 53; Conservative 29; Mismatches 62; Indels 29; Gaps 9;
QY 20 ITSVVLACSVTSEVHMIDNNGIKSIGVTFTDTDKGLQIKT--DLKGLPAGEHGFHIHE 77
Db 18 IKKLLSMAATKAVAVLKGDPVQ--GIINFQKESNGPVKWGSIKGLTEGLHGFHVHE 75
QY 78 GSSCGPAEDHGLTAGLAQH-GHYDPDKTGKHEGPLG-NGHKGDLPRLVVRADGTAKETL 135
Db 76 FGD-----NTAGCTSAGPHFNP-LSRKHGPKDEERHVGDLGNVTADKGVADVSI 125
QY 136 LAPRLTVK---EIKGRVTMHH-----AGDNYSDKPLPLGGGGARIACGVI 178
Db 126 EDSVISLSGDSHSIIIGRTLIVVHEKADDLKGKGNNESTK---TGNAGSRLACGVI 175
RESULT 27
AAR24235
ID AAR24235 standard; protein; 203 AA.
XX
AC AAR24235;
XX
XX 25-MAR-2003 (revised)
DT 25-NOV-1992 (first entry)
XX
DE GAG fusion protein with SOD according to a formula.
XX

```

```

KW Glycosamino:glycan; superoxidisedismutase; tissue damage;
KW autoimmune disease; rheumatoid arthritis; osteoarthritis; ss.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT /note= "GAG binding motif"
FT Peptide 32..203
FT /note= "SOD"
XX
PN WO9207935-A1.
XX
PD 14-MAY-1992.
XX
PF 01-NOV-1991; 91WO-US008105.
XX
PR 01-NOV-1990; 90US-00608539.
PR 02-NOV-1990; 90US-00608569.
XX
PA (SCRI ) SCRIPPS RES INST.
XX
PI Tainer JA, Kuhn L, Boissinot M, Fisher C, Parge HE, Griffin JH;
PI Mullenbach GT, Hallewell RA;
XX
DR WPI; 1992-183671/22.
XX
PS Fusion proteins with glycosamino:glycan-binding and superoxidisedismutase
PS activities - reduce tissue damage caused by super:oxide radicals, useful
PS in treating autoimmune diseases e.g. rheumatoid arthritis and
PS osteoarthritis.
XX
PS Claim 8; Fig 1; 140pp; English.
XX
CC The fusion protein comprising the a glycosaminoglycan binding region and
CC human superoxide dismutase, joined via a linker region was constructed
CC according to the formula M-(Z-M)3-SOD where Z is the peptide -EXTLRKWLK-.
CC The fusion protein is useful for extending the in vivo lifetimes of
CC biologically active cpds. such as SOD and for targetting them to specific
CC cell surfaces or substrates. The glycosaminoglycan (GAG) binding protein
CC is formed into a fusion protein with SOD to increase stability, plasma
CC half-life and ease of purification of SOD. SOD is useful for reduction of
CC tissue damage caused by oxygen radicals and is used in the treatment of
CC autoimmune diseases e.g. rheumatoid and osteo-arthritis. See also
CC AAR24225-35, AAR27932-51. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 203 AA;
Query Match 18.1%; Score 171.5; DB 2; Length 203;
Best Local Similarity 31.5%; Pred. No. 7, 8e-11;
Matches 53; Conservative 26; Mismatches 60; Indels 29; Gaps 9;
QY 25 LACSVTSEVHMIDNNGIKSIGVTFTDTDKGLQIKT--DLKGLPAGEHGFHIHGGSGG 82
Db 29 LKMAATKAVAVLKGDPVQ--GIINFQKESNGPVKWGSIKGLTEGLHGFHVHEFGD-- 84
QY 83 PAEDHGLTAGLAQH-GHYDPDKTGKHEGPLG-NGHKGDLPRLVVRADGTAKETLAPRL 140
Db 85 -----NTAGCTSAGPHFNP-LSRKHGPKDEERHVGDLGNVTADKGVADVSI 136
QY 141 TVK---EIKGRVTMHH-----AGDNYSDKPLPLGGGGARIACGVI 178
Db 137 SLSGDSHSIIIGRTLIVVHEKADDLKGKGNNESTK---TGNAGSRLACGVI 181
RESULT 28
AAR24233
ID AAR24233 standard; protein; 209 AA.
XX
AC AAR24233;
XX
XX 25-MAR-2003 (revised)
DT 25-MAR-2003 (revised)

```

DT 25-NOV-1992 (first entry)  
XX AC  
XX DE  
XX GAG fusion protein with SOD according to a formula.  
XX KW Glycosamino-glycan; superoxidisedismutase; tissue damage;  
XX auto-immune disease; rheumatoid arthritis; osteoarthritis; ss.  
XX OS Synthetic.  
XX FH Location/Qualifiers  
FT Key  
FT Peptide 1..37  
FT /note= "GAG binding motif"  
FT Peptide 38..209  
FT /note= "SOD "  
XX W09207935-A1.  
XX PD  
XX 14-MAY-1992.  
XX PF  
XX 01-NOV-1991; 91WO-US008105.  
XX XX  
XX 01-NOV-1990; 90US-00608539.  
XX PR  
XX 02-NOV-1990; 90US-00608569.  
XX XX  
XX (SCRI ) SCRIPPS RES INST.  
XX PA  
XX Tainer JA, Kuhn L, Boissinot M, Fisher C, Parge HE, Griffin JH;  
XX PI Mullenbach GT, Hallewell RA;  
XX PI  
XX WPI; 1992-183671/22.  
XX DR  
XX Fusion proteins with glycosamino-glycan-binding and superoxidisedismutase  
XX PT activities - reduce tissue damage caused by superoxide radicals, useful  
XX PT in treating autoimmune diseases e.g. rheumatoid arthritis and  
XX PT osteoarthritis.  
XX XX  
XX Claim 8; Fig 1; 140pp; English.  
XX XX  
XX The fusion protein comprising the a glycosaminoglycan binding region and  
XX human superoxide dismutase, joined via a linker region was constructed  
XX according to the formula M-(2-M)-SOD where Z is the peptide -YKIIKKLLLES  
XX -. The fusion protein is useful for extending the in vivo lifetimes of  
XX biologically active cpds. such as SOD and for targeting them to specific  
XX cell surfaces or substrates. The glycosaminoglycan (GAG) binding protein  
XX is formed into a fusion protein with SOD to increase stability, plasma  
XX half-life and ease of purification of SOD. SOD is useful for reduction of  
XX tissue damage caused by oxygen radicals and is used in the treatment of  
XX autoimmune diseases e.g. rheumatoid and osteo-arthritis. See also  
XX AAR24225-35, AAR27932-51. (Updated on 25-MAR-2003 to correct PN field.)  
XX CC  
XX (Updated on 25-MAR-2003 to correct PI field.)  
XX XX  
XX Sequence 209 AA;  
XX  
XX Query Match 18.1%; Score 171.5; DB 2; Length 209;  
XX Best Local Similarity 30.6%; Pred. No. 8.1e-11;  
XX Matches 53; Conservative 29; Mismatches 62; Indels 29; Gaps 9;  
XX  
XX 20 ITSVLACVTSVHMDNGIKOSIGTVTFDTDKGLQIKT--DLKGLPAGEGHGPHIHE 77  
XX 30 IKLLLESMAATKAVAVLKGGPQVQ--GIINFQKESNGPVKVMGSIKGLTEGLHGHVHE 87  
XX  
XX 78 GSGCGPAEHDGHLTALQAHG-HYDPDKTKGHEGPLG-NGHKGDLRLVVVKADGIAKETL 135  
XX 88 FGD-----NTAGCTAGPHNP-LSRKHGPGKDEERHVGDLGNVTADKGVADSVI 137  
XX  
XX 136 LAPRLTVK---ETKGRVTMIH-----AGDNYSDKPLPLGGGGARIACGVI 178  
XX 138 EDSVISLSGDHSIIGRTLTVVHEKADDLGKGNEESTK---TGNAGSLACGVI 187  
XX  
XX RESULT 29  
XX AAG11627  
XX ID AAG11627 standard; protein; 152 AA.

XX AAG11627;  
XX AC  
XX DT  
XX 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 10415.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX XX  
XX 06-SEP-2000.  
XX PD  
XX 25-FEB-2000; 2000EP-00301439.  
XX PF  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX PR 09-MAR-1999; 99US-0123548P.  
XX PR 23-MAR-1999; 99US-0125788P.  
XX PR 25-MAR-1999; 99US-0126264P.  
XX PR 29-MAR-1999; 99US-0126785P.  
XX PR 01-APR-1999; 99US-0127462P.  
XX PR 06-APR-1999; 99US-0128234P.  
XX PR 08-APR-1999; 99US-0128714P.  
XX PR 16-APR-1999; 99US-0129845P.  
XX PR 19-APR-1999; 99US-0130077P.  
XX PR 21-APR-1999; 99US-0130449P.  
XX PR 23-APR-1999; 99US-0130510P.  
XX PR 28-APR-1999; 99US-0130891P.  
XX PR 30-APR-1999; 99US-0131449P.  
XX PR 30-APR-1999; 99US-0132048P.  
XX PR 30-APR-1999; 99US-0132407P.  
XX PR 04-MAY-1999; 99US-0132484P.  
XX PR 05-MAY-1999; 99US-0132485P.  
XX PR 06-MAY-1999; 99US-0132486P.  
XX PR 06-MAY-1999; 99US-0132487P.  
XX PR 07-MAY-1999; 99US-0132863P.  
XX PR 11-MAY-1999; 99US-0134256P.  
XX PR 14-MAY-1999; 99US-0134218P.  
XX PR 14-MAY-1999; 99US-0134219P.  
XX PR 14-MAY-1999; 99US-0134221P.  
XX PR 14-MAY-1999; 99US-0134370P.  
XX PR 18-MAY-1999; 99US-0134768P.  
XX PR 19-MAY-1999; 99US-0134941P.  
XX PR 20-MAY-1999; 99US-0135124P.  
XX PR 21-MAY-1999; 99US-0135353P.  
XX PR 24-MAY-1999; 99US-0135629P.  
XX PR 25-MAY-1999; 99US-0136021P.  
XX PR 27-MAY-1999; 99US-0136392P.  
XX PR 28-MAY-1999; 99US-0136782P.  
XX PR 01-JUN-1999; 99US-0137222P.  
XX PR 03-JUN-1999; 99US-0137528P.  
XX PR 04-JUN-1999; 99US-0137502P.  
XX PR 07-JUN-1999; 99US-0137724P.  
XX PR 08-JUN-1999; 99US-0138094P.  
XX PR 10-JUN-1999; 99US-0138540P.  
XX PR 10-JUN-1999; 99US-0138847P.  
XX PR 14-JUN-1999; 99US-0139119P.  
XX PR 16-JUN-1999; 99US-0139452P.  
XX PR 16-JUN-1999; 99US-0139453P.  
XX PR 17-JUN-1999; 99US-0139492P.  
XX PR 18-JUN-1999; 99US-0139454P.  
XX PR 18-JUN-1999; 99US-0139455P.  
XX PR 18-JUN-1999; 99US-0139456P.  
XX PR 18-JUN-1999; 99US-0139457P.  
XX PR 18-JUN-1999; 99US-0139458P.  
XX PR 18-JUN-1999; 99US-0139459P.  
XX PR 18-JUN-1999; 99US-0139460P.  
XX PR 18-JUN-1999; 99US-0139461P.



ID AAG07549 standard; protein; 152 AA.  
XX AAG07549;  
AC  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4746.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 03-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 29-MAR-1999; 99US-0126264P.  
PR 01-APR-1999; 99US-0126785P.  
PR 08-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 16-APR-1999; 99US-0128714P.  
PR 19-APR-1999; 99US-0129845P.  
PR 21-APR-1999; 99US-0130077P.  
PR 23-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 21-JUN-1999;  
PR 22-JUN-1999;  
PR 23-JUN-1999;  
PR 23-JUN-1999;  
PR 24-JUN-1999;  
PR 24-JUN-1999;  
PR 28-JUN-1999;  
PR 29-JUN-1999;  
PR 30-JUN-1999;  
PR 01-JUL-1999;  
PR 01-JUL-1999;  
PR 02-JUL-1999;  
PR 06-JUL-1999;  
PR 08-JUL-1999;  
PR 09-JUL-1999;  
PR 12-JUL-1999;  
PR 13-JUL-1999;  
PR 14-JUL-1999;  
PR 15-JUL-1999;  
PR 16-JUL-1999;  
PR 16-JUL-1999;  
PR 19-JUL-1999;  
PR 19-JUL-1999;  
PR 19-JUL-1999;  
PR 19-JUL-1999;  
PR 19-JUL-1999;  
PR 20-JUL-1999;  
PR 20-JUL-1999;  
PR 21-JUL-1999;  
PR 21-JUL-1999;  
PR 22-JUL-1999;  
PR 22-JUL-1999;  
PR 22-JUL-1999;  
PR 23-JUL-1999;  
PR 23-JUL-1999;  
PR 26-JUL-1999;  
PR 27-JUL-1999;  
PR 27-JUL-1999;  
PR 27-JUL-1999;  
PR 28-JUL-1999;  
PR 02-AUG-1999;  
PR 02-AUG-1999;  
PR 03-AUG-1999;  
PR 04-AUG-1999;  
PR 04-AUG-1999;  
PR 05-AUG-1999;  
PR 06-AUG-1999;  
PR 06-AUG-1999;  
PR 09-AUG-1999;  
PR 09-AUG-1999;  
PR 10-AUG-1999;  
PR 11-AUG-1999;  
PR 12-AUG-1999;  
PR 13-AUG-1999;  
PR 13-AUG-1999;  
PR 16-AUG-1999;  
PR 17-AUG-1999;  
PR 18-AUG-1999;  
PR 20-AUG-1999;  
PR 20-AUG-1999;  
PR 23-AUG-1999;  
99US-0139461P.  
99US-0139462P.  
99US-0139463P.  
99US-0139750P.  
99US-0139763P.  
99US-0139817P.  
99US-0139899P.  
99US-0140353P.  
99US-0140354P.  
99US-0140695P.  
99US-0140823P.  
99US-0140991P.  
99US-0141287P.  
99US-0141842P.  
99US-0142154P.  
99US-0142055P.  
99US-0142390P.  
99US-0142803P.  
99US-0142920P.  
99US-0142977P.  
99US-0143542P.  
99US-0143624P.  
99US-0144005P.  
99US-0144085P.  
99US-0144086P.  
99US-0144325P.  
99US-0144331P.  
99US-0144332P.  
99US-0144333P.  
99US-0144334P.  
99US-0144335P.  
99US-0144352P.  
99US-0144632P.  
99US-0144884P.  
99US-0144814P.  
99US-0145086P.  
99US-0145088P.  
99US-0145085P.  
99US-0145087P.  
99US-0145089P.  
99US-0145192P.  
99US-0145145P.  
99US-0145218P.  
99US-0145224P.  
99US-0145276P.  
99US-0145913P.  
99US-0145918P.  
99US-0145919P.  
99US-0145951P.  
99US-0146386P.  
99US-0146388P.  
99US-0146389P.  
99US-0147038P.  
99US-0147204P.  
99US-0147302P.  
99US-0147192P.  
99US-0147260P.  
99US-0147303P.  
99US-0147416P.  
99US-0147493P.  
99US-0147935P.  
99US-0148171P.  
99US-0148319P.  
99US-0148341P.  
99US-0148565P.  
99US-0148684P.  
99US-0149368P.  
99US-0149175P.  
99US-0149426P.  
99US-0149722P.  
99US-0149723P.  
99US-0149929P.  
99US-0149902P.

|  |   |                |  |
|--|---|----------------|--|
| PR   | 23-AUG-1999;  | 99US-0149930P. |  |
| PR   | 25-AUG-1999;  | 99US-0150566P. |  |
| PR   | 26-AUG-1999;  | 99US-0150884P. |  |
| PR   | 27-AUG-1999;  | 99US-0151065P. |  |
| PR   | 27-AUG-1999;  | 99US-0151066P. |  |
| PR   | 27-AUG-1999;  | 99US-0151080P. |  |
| PR   | 30-AUG-1999;  | 99US-0151303P. |  |
| PR   | 31-AUG-1999;  | 99US-0151438P. |  |
| PR   | 01-SEP-1999;  | 99US-0151930P. |  |
| PR   | 07-SEP-1999;  | 99US-0152363P. |  |
| PR   | 10-SEP-1999;  | 99US-0153070P. |  |
| PR   | 13-SEP-1999;  | 99US-0153758P. |  |
| PR   | 15-SEP-1999;  | 99US-0154018P. |  |
| PR   | 28-SEP-1999;  | 99US-0156458P. |  |
| PR   | 29-SEP-1999;  | 99US-0156596P. |  |
| PR   | 04-OCT-1999;  | 99US-0157117P. |  |
| PR   | 05-OCT-1999;  | 99US-0157753P. |  |
| PR   | 06-OCT-1999;  | 99US-0157865P. |  |
| PR   | 07-OCT-1999;  | 99US-0158029P. |  |
| PR   | 08-OCT-1999;  | 99US-0158232P. |  |
| PR   | 12-OCT-1999;  | 99US-0158369P. |  |
| PR   | 13-OCT-1999;  | 99US-0159293P. |  |
| PR   | 13-OCT-1999;  | 99US-0159294P. |  |
| PR   | 13-OCT-1999;  | 99US-0159295P. |  |
| PR   | 14-OCT-1999;  | 99US-0159329P. |  |
| PR   | 14-OCT-1999;  | 99US-0159330P. |  |
| PR   | 14-OCT-1999;  | 99US-0159331P. |  |
| PR   | 14-OCT-1999;  | 99US-0159637P. |  |
| PR   | 18-OCT-1999;  | 99US-0159638P. |  |
| PR   | 18-OCT-1999;  | 99US-0159584P. |  |
| PR   | 21-OCT-1999;  | 99US-0160741P. |  |
| PR   | 21-OCT-1999;  | 99US-0160767P. |  |
| PR   | 21-OCT-1999;  | 99US-0160768P. |  |
| PR   | 21-OCT-1999;  | 99US-0160770P. |  |
| PR   | 21-OCT-1999;  | 99US-0160814P. |  |
| PR   | 21-OCT-1999;  | 99US-0160815P. |  |
| PR   | 22-OCT-1999;  | 99US-0160980P. |  |
| PR   | 22-OCT-1999;  | 99US-0160981P. |  |
| PR   | 22-OCT-1999;  | 99US-0160989P. |  |
| PR   | 25-OCT-1999;  | 99US-0161404P. |  |
| PR   | 25-OCT-1999;  | 99US-0161405P. |  |
| PR   | 25-OCT-1999;  | 99US-0161406P. |  |
| PR   | 26-OCT-1999;  | 99US-0161359P. |  |
| PR   | 26-OCT-1999;  | 99US-0161360P. |  |
| PR   | 26-OCT-1999;  | 99US-0161361P. |  |
| PR   | 28-OCT-1999;  | 99US-0161920P. |  |
| PR   | 28-OCT-1999;  | 99US-0161922P. |  |
| PR   | 28-OCT-1999;  | 99US-0161993P. |  |
| PR   | 29-OCT-1999;  | 99US-0162142P. |  |
| Query Match 18.0%; Score 171; DB 3; Length 152;                |   |                |  |
| Best Local Similarity 34.7%; Pred. NO. 5.9e-11;                |   |                |  |
| Matches 50; Conservative 16; Mismatches 56; Indels 22; Gaps 8; |   |                |  |
| Qy   | 46 GTVFTDIDKGL-QIKTDLKGIPAGEHGFHIEGSCGPAEDGHILTAGLQAHG-HYDPD 103                |                |  |
| Db   | 16 GTIIFTEQBGDVTTSVGTGVLKPLGHGFHVALGD-----TTNGCMSTGHEHNP 66                     |                |  |
| Qy   | 104 KTKG-HEGL- <del>CNGHKGDLPLRVKADGI</del> AKETLL---APRLTVKEIKGRVTMIHAGDGN 158 |                |  |
| Db   | 67 --GKTHGAPEDANRHAGDLGNITVGDGDTATFTIDCCIPLTGPNSIVGRAVVVHADPDD 124              |                |  |
| Qy   | 159 YS-----DKPLPLGGGARIACGVI 178  |                |  |
| Db   | 125 LGKGGHLSLATGNAGGRVACGII 148   |                |  |

|          |  |
|----------|--|
| AAG47395 |  |
| ID       | AAG47395 standard; protein; 152 AA.                                      |
| XX       |  |
| AC       | AAG47395;  |
| XX       |  |
| DT       | 18-OCT-2000 (first entry)  |
| XX       |  |
| DE       | Arabidopsis thaliana protein fragment SEQ ID NO: 59733.                  |
| XX       |  |
| KW       | Protein identification; signal transduction pathway; metabolic pathway;  |
| KW       | hybridisation assay; genetic mapping; gene expression control; promoter; |
| KW       | termination sequence.  |
| XX       |  |
| OS       | Arabidopsis thaliana.  |
| XX       |  |
| PN       | EP1033405-A2.  |
| XX       |  |
| PD       | 06-SEP-2000.   |
| XX       |  |
| PF       | 25-FEB-2000; 2000EP-00301439.  |
| XX       |  |
| PR       | 25-FEB-1999; 99US-0121825P.  |
| PR       | 05-MAR-1999; 99US-0123180P.  |
| PR       | 09-MAR-1999; 99US-0123548P.  |
| PR       | 23-MAR-1999; 99US-0125788P.  |
| PR       | 25-MAR-1999; 99US-0126264P.  |
| PR       | 29-MAR-1999; 99US-0126785P.  |
| PR       | 01-APR-1999; 99US-0127462P.  |
| PR       | 06-APR-1999; 99US-0128234P.  |
| PR       | 08-APR-1999; 99US-0128714P.  |
| PR       | 16-APR-1999; 99US-0129845P.  |
| PR       | 19-APR-1999; 99US-0130077P.  |
| PR       | 21-APR-1999; 99US-0130449P.  |
| PR       | 23-APR-1999; 99US-0130510P.  |
| PR       | 23-APR-1999; 99US-0130891P.  |
| PR       | 28-APR-1999; 99US-0131449P.  |
| PR       | 30-APR-1999; 99US-0132048P.  |
| PR       | 30-APR-1999; 99US-0132407P.  |
| PR       | 04-MAY-1999; 99US-0132484P.  |
| PR       | 05-MAY-1999; 99US-0132485P.  |
| PR       | 06-MAY-1999; 99US-0132486P.  |
| PR       | 06-MAY-1999; 99US-0132487P.  |
| PR       | 07-MAY-1999; 99US-0132863P.  |
| PR       | 11-MAY-1999; 99US-0134256P.  |
| PR       | 14-MAY-1999; 99US-0134218P.  |
| PR       | 14-MAY-1999; 99US-0134219P.  |
| PR       | 14-MAY-1999; 99US-0134221P.  |
| PR       | 14-MAY-1999; 99US-0134370P.  |
| PR       | 18-MAY-1999; 99US-0134768P.  |
| PR       | 19-MAY-1999; 99US-0134941P.  |
| PR       | 20-MAY-1999; 99US-0135124P.  |
| PR       | 21-MAY-1999; 99US-0135353P.  |
| PR       | 24-MAY-1999; 99US-0135629P.  |
| PR       | 25-MAY-1999; 99US-0136021P.  |
| PR       | 27-MAY-1999; 99US-0136392P.  |
| PR       | 28-MAY-1999; 99US-0136782P.  |
| PR       | 01-JUN-1999; 99US-0137222P.  |
| PR       | 03-JUN-1999; 99US-0137528P.  |
| PR       | 04-JUN-1999; 99US-0137502P.  |
| PR       | 07-JUN-1999; 99US-0137724P.  |
| PR       | 08-JUN-1999; 99US-0138094P.  |
| PR       | 10-JUN-1999; 99US-0138540P.  |
| PR       | 10-JUN-1999; 99US-0138847P.  |
| PR       | 14-JUN-1999; 99US-0139119P.  |
| PR       | 16-JUN-1999; 99US-0139452P.  |
| PR       | 16-JUN-1999; 99US-0139453P.  |
| PR       | 17-JUN-1999; 99US-0139432P.  |
| PR       | 18-JUN-1999; 99US-0139454P.  |
| PR       | 18-JUN-1999; 99US-0139455P.  |
| PR       | 18-JUN-1999; 99US-0139456P.  |
| PR       | 18-JUN-1999; 99US-0139457P.  |
| PR       | 18-JUN-1999; 99US-0139458P.  |
| PR       | 18-JUN-1999; 99US-0139459P.  |

|    |              |                |   |              |                |
|----|--------------|----------------|---|--------------|----------------|
| PR | 18-JUN-1999; | 99US-0139460P. | PR  | 23-AUG-1999; | 99US-0149902P. |
| PR | 18-JUN-1999; | 99US-0139461P. | PR  | 23-AUG-1999; | 99US-0149930P. |
| PR | 18-JUN-1999; | 99US-0139462P. | PR  | 25-AUG-1999; | 99US-0150566P. |
| PR | 18-JUN-1999; | 99US-0139463P. | PR  | 26-AUG-1999; | 99US-0150884P. |
| PR | 18-JUN-1999; | 99US-0139750P. | PR  | 27-AUG-1999; | 99US-0151065P. |
| PR | 18-JUN-1999; | 99US-0139763P. | PR  | 27-AUG-1999; | 99US-0151066P. |
| PR | 21-JUN-1999; | 99US-0139817P. | PR  | 27-AUG-1999; | 99US-0151080P. |
| PR | 22-JUN-1999; | 99US-0139899P. | PR  | 30-AUG-1999; | 99US-0151303P. |
| PR | 23-JUN-1999; | 99US-0140353P. | PR  | 31-AUG-1999; | 99US-0151438P. |
| PR | 23-JUN-1999; | 99US-0140354P. | PR  | 01-SEP-1999; | 99US-0151930P. |
| PR | 24-JUN-1999; | 99US-0140695P. | PR  | 07-SEP-1999; | 99US-0152363P. |
| PR | 28-JUN-1999; | 99US-0140823P. | PR  | 10-SEP-1999; | 99US-0153070P. |
| PR | 29-JUN-1999; | 99US-0140991P. | PR  | 13-SEP-1999; | 99US-0153758P. |
| PR | 30-JUN-1999; | 99US-0141287P. | PR  | 15-SEP-1999; | 99US-0154018P. |
| PR | 01-JUL-1999; | 99US-0141842P. | PR  | 16-SEP-1999; | 99US-0154039P. |
| PR | 01-JUL-1999; | 99US-0142154P. | PR  | 15-SEP-1999; | 99US-0154779P. |
| PR | 02-JUL-1999; | 99US-0142055P. | PR  | 22-SEP-1999; | 99US-0155139P. |
| PR | 06-JUL-1999; | 99US-0142390P. | PR  | 23-SEP-1999; | 99US-0155486P. |
| PR | 08-JUL-1999; | 99US-0142803P. | PR  | 24-SEP-1999; | 99US-0155659P. |
| PR | 09-JUL-1999; | 99US-0142920P. | PR  | 28-SEP-1999; | 99US-0156458P. |
| PR | 12-JUL-1999; | 99US-0142977P. | PR  | 29-SEP-1999; | 99US-0156596P. |
| PR | 13-JUL-1999; | 99US-0143542P. | PR  | 04-OCT-1999; | 99US-0157117P. |
| PR | 14-JUL-1999; | 99US-0143624P. | PR  | 05-OCT-1999; | 99US-0157753P. |
| PR | 15-JUL-1999; | 99US-0144005P. | PR  | 06-OCT-1999; | 99US-0157865P. |
| PR | 16-JUL-1999; | 99US-0144085P. | PR  | 07-OCT-1999; | 99US-0158029P. |
| PR | 16-JUL-1999; | 99US-0144086P. | PR  | 08-OCT-1999; | 99US-0158232P. |
| PR | 19-JUL-1999; | 99US-0144325P. | PR  | 12-OCT-1999; | 99US-0158369P. |
| PR | 19-JUL-1999; | 99US-0144331P. | PR  | 13-OCT-1999; | 99US-0159293P. |
| PR | 19-JUL-1999; | 99US-0144332P. | PR  | 13-OCT-1999; | 99US-0159294P. |
| PR | 19-JUL-1999; | 99US-0144333P. | PR  | 13-OCT-1999; | 99US-0159295P. |
| PR | 19-JUL-1999; | 99US-0144334P. | PR  | 14-OCT-1999; | 99US-0159329P. |
| PR | 20-JUL-1999; | 99US-0144335P. | PR  | 14-OCT-1999; | 99US-0159330P. |
| PR | 20-JUL-1999; | 99US-0144352P. | PR  | 14-OCT-1999; | 99US-0159331P. |
| PR | 20-JUL-1999; | 99US-0144632P. | PR  | 14-OCT-1999; | 99US-0159637P. |
| PR | 20-JUL-1999; | 99US-0144884P. | PR  | 14-OCT-1999; | 99US-0159638P. |
| PR | 21-JUL-1999; | 99US-0144814P. | PR  | 18-OCT-1999; | 99US-0159584P. |
| PR | 21-JUL-1999; | 99US-0145086P. | PR  | 21-OCT-1999; | 99US-0160741P. |
| PR | 21-JUL-1999; | 99US-0145088P. | PR  | 21-OCT-1999; | 99US-0160767P. |
| PR | 22-JUL-1999; | 99US-0145085P. | PR  | 21-OCT-1999; | 99US-0160768P. |
| PR | 22-JUL-1999; | 99US-0145087P. | PR  | 21-OCT-1999; | 99US-0160770P. |
| PR | 22-JUL-1999; | 99US-0145089P. | PR  | 21-OCT-1999; | 99US-0160814P. |
| PR | 22-JUL-1999; | 99US-0145192P. | PR  | 21-OCT-1999; | 99US-0160815P. |
| PR | 23-JUL-1999; | 99US-0145145P. | PR  | 22-OCT-1999; | 99US-0160980P. |
| PR | 23-JUL-1999; | 99US-0145218P. | PR  | 22-OCT-1999; | 99US-0160981P. |
| PR | 23-JUL-1999; | 99US-0145224P. | PR  | 22-OCT-1999; | 99US-0160989P. |
| PR | 26-JUL-1999; | 99US-0145376P. | PR  | 25-OCT-1999; | 99US-0161404P. |
| PR | 27-JUL-1999; | 99US-0145913P. | PR  | 25-OCT-1999; | 99US-0161405P. |
| PR | 27-JUL-1999; | 99US-0145918P. | PR  | 25-OCT-1999; | 99US-0161406P. |
| PR | 27-JUL-1999; | 99US-0145919P. | PR  | 25-OCT-1999; | 99US-0161359P. |
| PR | 28-JUL-1999; | 99US-0145951P. | PR  | 26-OCT-1999; | 99US-0161360P. |
| PR | 02-AUG-1999; | 99US-0146386P. | PR  | 26-OCT-1999; | 99US-0161361P. |
| PR | 02-AUG-1999; | 99US-0146388P. | PR  | 28-OCT-1999; | 99US-0161920P. |
| PR | 02-AUG-1999; | 99US-0146389P. | PR  | 28-OCT-1999; | 99US-0161992P. |
| PR | 03-AUG-1999; | 99US-0147038P. | PR  | 28-OCT-1999; | 99US-0161993P. |
| PR | 04-AUG-1999; | 99US-0147204P. | PR  | 29-OCT-1999; | 99US-0162142P. |
| PR | 05-AUG-1999; | 99US-0147302P. | Query Match 18.0%; Score 171; DB 3; Length 152;                         |              |                |
| PR | 05-AUG-1999; | 99US-0147192P. | Best Local Similarity 34.7%; Pred.No.5.9e-11;                           |              |                |
| PR | 05-AUG-1999; | 99US-0147260P. | Matches 50; Conservative 16; Mismatches 56; Indels 22; Gaps 8;          |              |                |
| PR | 06-AUG-1999; | 99US-0147303P. | QY 46 GTVFTDTKGL-QIKTDLKLPGFHEGSGCPAEHDGHLTAGLQAHG-HYDPP 103            |              |                |
| PR | 06-AUG-1999; | 99US-0147416P. | Db 16 GTIFFTEGDGVTTSVSGTSGLKPGLHGFHVHALGD-----TTNGCMSTGPHFNPD 66        |              |                |
| PR | 09-AUG-1999; | 99US-0147493P. | QY 104 KTGK-HEGFL-GNGHKGDLPRLVVKADGIKAKETLL---APRLTVKGIKRTVMIHAGGDN 158 |              |                |
| PR | 10-AUG-1999; | 99US-0148171P. | Db 67 --GKTHGAPEDANRHAGDLGNITVGDDGTATFTITDCQIPLTGPNSIVGRAVVVHADPPD 124  |              |                |
| PR | 11-AUG-1999; | 99US-0148319P. | QY 159 YS----DKPLPLGGGGARIACGVI 178                                     |              |                |
| PR | 12-AUG-1999; | 99US-0148341P. | Db 125 LGKGHELSLATGNAGGRVACGII 148                                      |              |                |
| PR | 13-AUG-1999; | 99US-0148565P. |   |              |                |
| PR | 16-AUG-1999; | 99US-0149368P. |   |              |                |
| PR | 17-AUG-1999; | 99US-0149175P. |   |              |                |
| PR | 18-AUG-1999; | 99US-0149426P. |   |              |                |
| PR | 20-AUG-1999; | 99US-0149722P. |   |              |                |
| PR | 20-AUG-1999; | 99US-0149723P. |   |              |                |
| PR | 20-AUG-1999; | 99US-0149929P. |   |              |                |

```
RESULT 32
AAG31513
ID AAG31513 standard; protein; 152 AA.
XX AC AAG31513;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 37856.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127463P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130691P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132566P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137503P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140693P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140911P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
```



```
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159388P.
PR 14-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

```
Query Match 18.0%; Score 171; DB 3; Length 152;
Best Local Similarity 34.7%; Pred. No. 5.9e-11;
Matches 50; Conservative 16; Mismatches 56; Indels 22; Gaps 8;
QY 46 GTVTFTDQKL-QIKTDLKGPLAGHGFIHHEGSGCGPAEHGHLTAGLQAHG-HYDPD 103
Db 16 GTTFTQEGDGVTVTSVSGLKPGLGHFVHALGD-----TTNGCMSTGPHFND 66
QY 104 KTGK-HEGPL-GNGHKGDPLRLVVKADGIKAKETLL---APRLVFKIKGTVMIHAGGN 158
Db 67 --GKTHGAPDANRHAGDLGNITVGDDGTATFTITDQIPLTGPNISVGRVVVHADPDD 124
QY 159 YS----DKPLPLGGGGARTACGVI 178
Db 125 LGKGHSLSLATGNAGGRVACGII 148
```

```
RESULT 33
AAG31512
ID AAG31512 standard; protein; 185 AA.
XX
AC AAG31512;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37855.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 23-MAR-1999; 99US-0123548P.
PR 25-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 27-MAY-1999; 99US-0136782P.
PR 28-MAY-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
```



|  |     |  |    |                           |    |
|--|-----|--|----|---------------------------|----|
| RESULT 34                              | AC  | AAG06932;  | XX | 17-OCT-2000 (first entry) | XX |
| AAP81018                               | DT  |  | XX |                           | XX |
| ID AAP81018 standard; protein; 174 AA. | DE  | Arabidopsis thaliana protein fragment SEQ ID NO: 3892.                   | XX |                           | XX |
| XX                                     | XX  |  | XX |                           | XX |
| AC                                     | XX  | Protein identification; signal transduction pathway; metabolic pathway;  | XX |                           | XX |
| 10-MAR-2003 (revised)                  | KW  | hybridisation assay; genetic mapping; gene expression control; promoter; | XX |                           | XX |
| 01-NOV-1990 (first entry)              | KW  | termination sequence.  | XX |                           | XX |
| XX                                     | XX  |  | XX |                           | XX |
| DE                                     | OS  | Arabidopsis thaliana.  | XX |                           | XX |
| XX                                     | XX  |  | XX |                           | XX |
| XX                                     | PN  | EP1033405-A2.  | XX |                           | XX |
| XX                                     | PD  | 06-SEP-2000.   | XX |                           | XX |
| XX                                     | XX  |  | XX |                           | XX |
| XX                                     | PF  | 25-FEB-2000; 2000EP-00301439.  | XX |                           | XX |
| XX                                     | XX  |  | XX |                           | XX |
| XX                                     | PR  | 25-FEB-1999; 99US-0121825P.  | XX |                           | XX |
| XX                                     | PR  | 05-MAR-1999; 99US-0123180P.  | XX |                           | XX |
| XX                                     | PR  | 09-MAR-1999; 99US-0123548P.  | XX |                           | XX |
| XX                                     | PR  | 23-MAR-1999; 99US-0125788P.  | XX |                           | XX |
| XX                                     | PR  | 25-MAR-1999; 99US-0126264P.  | XX |                           | XX |
| XX                                     | PR  | 29-MAR-1999; 99US-0126785P.  | XX |                           | XX |
| XX                                     | PR  | 01-APR-1999; 99US-0127462P.  | XX |                           | XX |
| XX                                     | PR  | 06-APR-1999; 99US-0128234P.  | XX |                           | XX |
| XX                                     | PR  | 08-APR-1999; 99US-0128714P.  | XX |                           | XX |
| XX                                     | PR  | 16-APR-1999; 99US-0129845P.  | XX |                           | XX |
| XX                                     | PR  | 19-APR-1999; 99US-0130077P.  | XX |                           | XX |
| XX                                     | PR  | 21-APR-1999; 99US-0130449P.  | XX |                           | XX |
| XX                                     | PR  | 23-APR-1999; 99US-0130510P.  | XX |                           | XX |
| XX                                     | PR  | 28-APR-1999; 99US-0130891P.  | XX |                           | XX |
| XX                                     | PR  | 30-APR-1999; 99US-0131449P.  | XX |                           | XX |
| XX                                     | PR  | 30-APR-1999; 99US-0132048P.  | XX |                           | XX |
| XX                                     | PR  | 04-MAY-1999; 99US-0132407P.  | XX |                           | XX |
| XX                                     | PR  | 05-MAY-1999; 99US-0132484P.  | XX |                           | XX |
| XX                                     | PR  | 06-MAY-1999; 99US-0132485P.  | XX |                           | XX |
| XX                                     | PR  | 06-MAY-1999; 99US-0132486P.  | XX |                           | XX |
| XX                                     | PR  | 07-MAY-1999; 99US-0132487P.  | XX |                           | XX |
| XX                                     | PR  | 11-MAY-1999; 99US-0132863P.  | XX |                           | XX |
| XX                                     | PR  | 14-MAY-1999; 99US-0134218P.  | XX |                           | XX |
| XX                                     | PR  | 14-MAY-1999; 99US-0134219P.  | XX |                           | XX |
| XX                                     | PR  | 14-MAY-1999; 99US-0134221P.  | XX |                           | XX |
| XX                                     | PR  | 14-MAY-1999; 99US-0134370P.  | XX |                           | XX |
| XX                                     | PR  | 18-MAY-1999; 99US-0134768P.  | XX |                           | XX |
| XX                                     | PR  | 19-MAY-1999; 99US-0134941P.  | XX |                           | XX |
| XX                                     | PR  | 20-MAY-1999; 99US-0135124P.  | XX |                           | XX |
| XX                                     | PR  | 21-MAY-1999; 99US-0135353P.  | XX |                           | XX |
| XX                                     | PR  | 24-MAY-1999; 99US-0135629P.  | XX |                           | XX |
| XX                                     | PR  | 25-MAY-1999; 99US-0136021P.  | XX |                           | XX |
| XX                                     | PR  | 27-MAY-1999; 99US-0136392P.  | XX |                           | XX |
| XX                                     | PR  | 28-MAY-1999; 99US-0136782P.  | XX |                           | XX |
| XX                                     | PR  | 01-JUN-1999; 99US-0137222P.  | XX |                           | XX |
| XX                                     | PR  | 03-JUN-1999; 99US-0137528P.  | XX |                           | XX |
| XX                                     | PR  | 04-JUN-1999; 99US-0137502P.  | XX |                           | XX |
| XX                                     | PR  | 07-JUN-1999; 99US-0137724P.  | XX |                           | XX |
| XX                                     | PR  | 08-JUN-1999; 99US-0138094P.  | XX |                           | XX |
| XX                                     | PR  | 10-JUN-1999; 99US-0138540P.  | XX |                           | XX |
| XX                                     | PR  | 10-JUN-1999; 99US-0138847P.  | XX |                           | XX |
| XX                                     | PR  | 14-JUN-1999; 99US-0139119P.  | XX |                           | XX |
| XX                                     | PR  | 16-JUN-1999; 99US-0139452P.  | XX |                           | XX |
| XX                                     | PR  | 16-JUN-1999; 99US-0139453P.  | XX |                           | XX |
| XX                                     | PR  | 17-JUN-1999; 99US-0139492P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139454P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139455P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139456P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139457P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139458P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139459P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139460P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139461P.  | XX |                           | XX |
| XX                                     | PR  | 18-JUN-1999; 99US-0139462P.  | XX |                           | XX |
| Query Match                            |     | 17.9%; Score 170.5; DB 1; Length 174;                                    |    |                           |    |
| Best Local Similarity                  |     | 31.0%; Pred. No. 8.1e-11;  |    |                           |    |
| Matches                                |     | Conservative 28; Mismatches 68; Indels 31; Gaps 10;                      |    |                           |    |
| QY                                     | 9   | TSVTISLTSITSVLACSVTSEVHMIDNDNGIKQIGITVTFDTDKGLQIKT--DLKGL 66             |    |                           |    |
| Db                                     | 4   | TATAIAVALAGFATV--AQAATKAVCVLKGDPVQ--GIINFQKESNGPFVWGSIRGL 59             |    |                           |    |
| QY                                     | 67  | PAGEHGFPHIEGSCGPAEDHGLTAGLQAHG--HYDPDKTKRHEGPLG--NGHKGDLPRLV 124         |    |                           |    |
| Db                                     | 60  | TEGLHGFHVEFGD-----NTAGCTSAGHFNP--LSRKHGPGXDERHVGDLQNT 109                |    |                           |    |
| QY                                     | 125 | VKADGIKAKETLLAPRITVKE---IKGRVTMHI-----AGGDNYSDKPLPLGGGGARIA 174          |    |                           |    |
| Db                                     | 110 | ADKDGADVDSIEDSVISLSDHDCIIGRTLVLVHEKADDLKGGNEESTK---TGNAGSRLA 166         |    |                           |    |
| QY                                     | 175 | CGVI 178   |    |                           |    |
| Db                                     | 167 | CGVI 170   |    |                           |    |
| RESULT 35                              |     |  |    |                           |    |
| AAG06932                               |     |  |    |                           |    |
| ID AAG06932 standard; protein; 152 AA. |     |  |    |                           |    |
| XX                                     |     |  |    |                           |    |



```
XX ABB59872;
AC
XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 6408.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
PR
XX (PEKE ) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX N-PSDB; ABL03975.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 6408; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO.int/pub/pub/published_pct_sequences
XX
XX Sequence 264 AA;
SQ
Query Match 17.9%; Score 170; DB 4; Length 264;
Best Local Similarity 32.2%; Pred. No. 1.7e-10;
Matches 57; Conservative 22; Mismatches 72; Indels 26; Gaps 8;
QY 18 TSITSVVLACSVTSEVHMDDNGI---KQSI-GTVTFT---DTDKGLQIKTKLGLPAG 69
Db 60 TGRAVLISGFGGSAVALINTTSVVDKTPIQGVVFTTTITADKKPEVVVDGVDGLSPG 119
QY 70 EHGPHIHEGSCGPAEHDGHLTAGLAQH-GHYDPDKTGHEGLNG---HKGDPLRLVV 125
Db 120 LHGLIHESG-----TSAGCSSVGEHYNP-RQSPHGSFAAGAEERHAGDLGNIRA 169
QY 126 KADGIAKETLLAPRLTVKEIKGTVMIHAGDNY-----SDKPLPLGGGARIACGVI 178
Db 170 DENGRAFTRFVDPVLEWIDIIGRAVILTANADDLGRGNDQSLTDGNSGRIACGII 226
RESULT 37
AAW17901
ID AAW17901 standard; protein; 154 AA.
XX
XX AAW17901;
XX
XX 28-JUL-1997 (first entry)
DT
XX
XX Human superoxide dismutase encoded by plasmid pYLBC-A/G-SOD.
XX
```

```
XX
KW Human; superoxide dismutase; hSOD; copper; zinc; Cu, Zn-SOD; yeast;
KW Saccharomyces cerevisiae; recombinant production.
XX
XX Homo sapiens.
OS
XX KR9209502-B.
PN
XX 17-OCT-1992.
PD
XX
XX 30-JUN-1990; 90KR-00009926.
PF
XX
XX 30-JUN-1990; 90KR-00009926.
PR
XX (LUCK-) LUCKY CO.
PA
XX
XX Cho J, Yang J;
PI
XX WPI; 1993-292994/37.
DR
XX N-PSDB; AAT65097.
DR
XX
XX New expression vector of human superoxidodismutase prepn. - by treating
PT M12mp 18-SOD contg. human SOD gene with restriction enzyme, etc.
PT
XX
XX Claim 2; Fig 1; 6pp; Korean.
XX
XX The new human superoxide dismutase (hSOD) expression vector designated
CC pYLBC-A/G-SOD was constructed by ligating together three fragments as
CC follows: (1) the NcoI-SalI fragment of plasmid M13mpl8-SOD (E. coli LUCK-
CC SOD-1E; KfCC 10699) containing the hSOD gene and coding for a polypeptide
CC of the present sequence; (2) the PstI-NcoI fragment from Saccharomyces
CC cerevisiae plasmid pYLBC-A/G-HGH (KfCC 10669) containing the A/G promoter
CC and (3) the PstI-SalI fragment of pYLBC-A/G-HGH containing the GAP
CC terminator
XX
XX Sequence 154 AA;
SQ
Query Match 17.8%; Score 169.5; DB 2; Length 154;
Best Local Similarity 32.5%; Pred. No. 8.9e-11;
Matches 53; Conservative 24; Mismatches 57; Indels 29; Gaps 9;
QY 30 TSEVHMDDNGIKQSIGTGTFTTDTDKGLQIKT--DLKGLPAGHGHPIHEGSCGPAEHD 87
Db 3 TRAVCVLKGDGPVQ--GIINFEOKESNGPVKWGSIKGLTEGLGHFVHEFGD----- 53
QY 88 GHLTAGLQAHG-HYDPDKTGHEGLG-NGHKGDPLRLVVKADGIAKETLLAPRLTVKE- 144
Db 54 --NTAGCTAGPHFNPILT-KHGGPKDEERHVGDLGNVTADKGDVADVSDSVLSLGD 110
QY 145 --IKGRTVMIH-----AGGQNYSDKPLPLGGGARIACGVI 178
Db 111 HCLIGRTLTVVHEKADDDLKGKNEESTK---TGNAGSRLACGVI 150
RESULT 38
AAAR27951
ID AAR27951 standard; protein; 183 AA.
XX
XX AAR27951;
XX
XX 25-MAR-2003 (revised)
DT
XX 25-NOV-1992 (first entry)
DT
XX
XX GAG fusion protein with SOD according to a formula.
XX
XX Glycosamino:glycan; superoxidodismutase; tissue damage;
KW autoimmune disease; rheumatoid arthritis; osteoarthritis; ss.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..172
FT /note= "SOD"
```



Search completed: October 26, 2004, 09:41:35  
Job time : 158 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 1, 2004, 17:08:50 ; Search time 2547 Seconds  
(without alignments)

3342.028 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 950

Sequence: 1 MKIKLFFVTSIVTISLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DRV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10009916/runat\_26102004.100223.1967/app\_query.fasta\_1.327  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10009916 @CGN 1.1 2527 @runat\_26102004.100223.1967 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DRV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.btg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| C 1        | 476.5 | 50.2        | 14203  | 1  | AE015188 Shigella    |
| C 2        | 476.5 | 50.2        | 290628 | 1  | AE016983 Shigella    |
| C 3        | 474.5 | 49.9        | 919    | 1  | ECU51242             |
| C 4        | 474.5 | 49.9        | 12347  | 1  | AE005387 Escherichia |

|      |       |      |        |   |           |
|------|-------|------|--------|---|-----------|
| C 5  | 474.5 | 49.9 | 110000 | 1 | U00096.17 |
| C 6  | 474.5 | 49.9 | 257071 | 1 | AP002558  |
| C 7  | 473.5 | 49.8 | 300475 | 1 | AE016761  |
| C 8  | 470.5 | 49.5 | 468    | 1 | ECSDC     |
| C 9  | 452   | 47.6 | 786    | 1 | AJ620904  |
| C 10 | 451   | 47.5 | 1274   | 1 | AF056931  |
| C 11 | 451   | 47.5 | 21913  | 1 | AE008762  |
| C 12 | 449   | 47.3 | 233050 | 1 | AL627271  |
| C 13 | 449   | 47.3 | 300523 | 1 | AE016838  |
| C 14 | 435.5 | 45.8 | 311249 | 1 | AE016860  |
| C 15 | 430   | 45.3 | 642    | 6 | AR319374  |
| C 16 | 429.5 | 44.2 | 435    | 6 | AR385611  |
| C 17 | 426   | 44.8 | 696    | 1 | STCZSDC   |
| C 18 | 426   | 44.8 | 840    | 1 | AJ620903  |
| C 19 | 426   | 44.8 | 62027  | 1 | AE008743  |
| C 20 | 421   | 44.3 | 10086  | 1 | AE014566  |
| C 21 | 421   | 44.3 | 10738  | 1 | AE009694  |
| C 22 | 420   | 44.2 | 3333   | 1 | AF007380  |
| C 23 | 417.5 | 43.9 | 110000 | 1 | AE004235  |
| C 24 | 417.5 | 43.9 | 12372  | 1 | AE005367  |
| C 25 | 414   | 43.6 | 11274  | 1 | AE005440  |
| C 26 | 414   | 43.6 | 34063  | 6 | AR204200  |
| C 27 | 414   | 43.6 | 49650  | 6 | BD184769  |
| C 28 | 413.5 | 43.5 | 718    | 7 | AF254764  |
| C 29 | 413.5 | 43.5 | 65219  | 1 | AE008737  |
| C 30 | 413.5 | 43.5 | 9799   | 1 | AE005440  |
| C 31 | 413   | 43.5 | 39824  | 6 | BD184776  |
| C 32 | 413   | 43.5 | 48908  | 6 | AR204241  |
| C 33 | 413   | 43.5 | 291136 | 1 | AP002556  |
| C 34 | 413   | 43.3 | 253750 | 1 | AP005345  |
| C 35 | 411   | 43.2 | 335862 | 1 | AP005089  |
| C 36 | 410   | 42.8 | 336182 | 1 | AE016813  |
| C 37 | 407   | 42.8 | 336182 | 1 | AE016813  |
| C 38 | 404.5 | 42.6 | 11361  | 1 | AE013684  |
| C 39 | 404   | 42.5 | 220050 | 1 | AJ414156  |
| C 40 | 404   | 42.5 | 290803 | 1 | AE017128  |
| C 41 | 399   | 42.0 | 348934 | 1 | EX640417  |
| C 42 | 394.5 | 41.5 | 971    | 1 | AB105404  |
| C 43 | 391   | 41.2 | 739    | 1 | HEACZSODB |
| C 44 | 391   | 41.2 | 841    | 1 | PHRSOPL   |
| C 45 | 391   | 41.2 | 348525 | 1 | EX640428  |

#### ALIGNMENTS

| RESULT 1   | LOCUS                         | DEFINITION   | ACCESSION | VERSION  | KEYWORDS | SOURCE     | ORGANISM    | REFERENCE | AUTHORS | TITLE  | JOURNAL | PUBMED | REFERENCE | AUTHORS |
|------------|-------------------------------|--|-----------|----------|----------|------------|-------------|-----------|---------|--|---------|--------|-----------|---------|
| AE015188/c | Shigella flexneri 2a str. 301 | 14203 bp DNA linear BCT 18-OCT-2002 of 412 of the complete genome. | AE015188  | AE005674 |          | AE015188.1 | GI:24052002 |           |         | Shigella flexneri 2a str. 301  |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | Shigella flexneri 2a str. 301  |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Shigella.  |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | 1 (bases 1 to 14203)   |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | Jin, Q., Yuan, Z. H., Xu, J. G., Wang, Y., Shen, Y., Lu, W. C., Wang, J. H., Liu, H., Yang, J., Yang, F., Qu, D., Zhang, X. B., Zhang, J. Y., Yang, G. W., Wu, H. T., Dong, J., Sun, L. L., Xue, Y., Zhao, A. L., Gao, Y. S., Zhu, J. P., Kan, B., Chen, S. X., Yao, Z. J., He, B. K., Chen, R. S., Ma, D. L., Qiang, B. Q., Wen, Y. M., Hou, Y. D. and Yu, J. |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157  |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | Nucleic Acids Res. 30 (20), 4432-4441 (2002)   |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | 2 (bases 1 to 14203)   |         |        |           |         |
|            |                               |  |           |          |          |            |             |           |         | Jin, Q., Shen, Y., Wang, J. H., Liu, H., Yang, J., Yang, F., Zhang, X. B., Zhang, J. Y., Yang, G. W., Wu, H. T., Dong, J., Sun, L. L., Xue, Y., Zhao, A. L., Gao, Y. S., Zhu, J. P., Chen, S. X., Yao, Z. J., Wang, Y., Lu, W. C., Qiang, B. Q., Wen, Y. M. and Hou, Y. D.   |         |        |           |         |

TITLE Direct Submission  
JOURNAL Submitted (08-MAY-2001) Microbial Genome Center of Chinese Ministry of Public Health, 100 YingXin Jie, XuanWu Qu, Beijing 100052, P.R. China

FEATURES  
source

Location/Qualifiers

1. .14203  
/organism="Shigella flexneri 2a str. 301"  
/mol\_type="genomic DNA"  
/strain="301"  
/serotype="2a"  
/db\_xref="taxon:198214"

## gene

128. .364  
/gene="ydhH"  
/locus\_tag="SF1670"

## CDS

128. .364  
/gene="ydhH"

/locus\_tag="SF1670"  
/note="Residues 1 to 78 of 78 are 98 pct identical to residues 1 to 78 of a 78 aa protein from Escherichia coli O157:H7 EDL933 ref: NP\_288079.1"

/codon\_start=1  
/transl\_table=11

/product="orf, conserved hypothetical protein"

/protein\_id="AAN43252.1"  
/db\_xref="GI:24052003"

/translation="MKFMNATGLPLQDLVFCASVYFPFFKAFAGFVIMLVVHRLR  
RGWIYAGDIWHPLMLDLSLFAICVLAILIAW"

## gene

358. .1224  
/gene="ydhJ"

/locus\_tag="SF1671"

## CDS

358. .1224  
/gene="ydhJ"

/locus\_tag="SF1671"  
/note="Residues 1 to 288 of 288 are 98 pct identical to residues 12 to 299 of a 299 aa protein from Escherichia coli O157:H7 EDL933 ref: NP\_288080.1"

/codon\_start=1  
/transl\_table=11

/product="putative membrane protein"

/protein\_id="AAN43253.1"  
/db\_xref="GI:24052004"

/translation="MVTMSIKTIKVFSTIIIVAVLAGWLWNYMOSPWTRGKIR  
AEQVSTPQSGRIVELNIKONQLVNAAGELLTIDKTFQIAELNAQALAKAQSDDLA  
KANNEANRRHLSQNFIFABELDTANLVKAAQASVNAQAATLKQTOWQLAQTEIRAP  
VSGWVNTNTRIGYADGKPLFALVDGSHSVFYGYFETKLRIHREGAPAQITLYSD  
NKTLOQHVSSIGRAIYDQSVESDLSLPDVKPNVPWVRLAQIVRPALDKVPGDVTL  
VSGTTCSTAVGQ"

1224. .3232  
/locus\_tag="SF1672"

/pseudo

## gene

1224. .3232  
/locus\_tag="SF1672"

/note="similar to Escherichia coli K12 orf, conserved hypothetical protein gi: 1787933 (671 aa). BLAST with identity of 98% in 670 aa. This CDS contains an in-frame stop codon. The sequence has been checked and is believed to be correct."

/pseudo

/codon\_start=1  
/transl\_table=11

/complement(3237. .3809)

## gene

/gene="sodC"

## CDS

/complement(3237. .3809)

/gene="sodC"

/locus\_tag="SF1673"  
/note="Residues 18 to 190 of 190 are 99 pct identical to residues 1 to 173 of a 173 aa protein from Escherichia coli O157:H7 EDL933 ref: NP\_288082.1"

/codon\_start=1  
/transl\_table=11

/product="superoxide dismutase precursor (Cu-Zn)"

/protein\_id="AAN43254.1"

## gene

/locus\_tag="SF1674"

## CDS

/complement(3839. .4735)

## CDS

/gene="ydhP"

## CDS

/locus\_tag="SF1674"

## CDS

/note="Residues 1 to 298 of 298 are 98 pct identical to residues 1 to 298 of a 298 aa protein from Escherichia coli K12 ref: NP\_416164.1"

/codon\_start=1  
/transl\_table=11

## CDS

/product="orf, conserved hypothetical protein"

/protein\_id="AAN43255.1"

## CDS

/db\_xref="GI:24052006"

## CDS

/translation="MVQRITIAPEGPEFSRFVMGYWRLMWNMSARQLVSTFEEHLDL  
GVTVDHADIYGGYQCEAAFGALKAPHLRMEIVSKCGIATTAARENVIGHYITD  
RNHIKGAESQLINLATDHLDLLIHRPDLMDADAVADAFKHLHOSKVRHFGVSNF  
TPAQFALLOSRLPFTLATNOVEISPVHQPLLDLTDLQOLRVPRMAWSCGGRRLF  
NDYFQRLDELAVVAEELNAGSIEQVVIWVLRPLPSQPLFIIGSKTERVRAAVEAE  
TLKMTROQWFRIRKAAALGYDVP"

/complement(4784. .5194)

## gene

/gene="ydhL"

## CDS

/locus\_tag="SF1675"

## CDS

/complement(4784. .5194)

## CDS

/gene="ydhL"

## CDS

/locus\_tag="SF1675"

## CDS

/note="Residues 12 to 136 of 136 are 100 pct identical to residues 1 to 125 of a 125 aa protein from Escherichia coli K12 ref: NP\_416165.1"

/codon\_start=1  
/transl\_table=11

## CDS

/product="orf, conserved hypothetical protein"

/protein\_id="AAN43256.1"

## CDS

/db\_xref="GI:24052007"

## CDS

/translation="MQAKIARROBMFASIMFGVHAECSRPGVLVQAKILLWTLT  
ASIMVILISEGEVAEQLEPPQSPCGICQSDRGFCRCFSRDRFRWNKMSDG  
EKQEVLRICRORLMKLRANKPASDDEFEQSLP"

5210. .5725  
/gene="ydhM"

## CDS

/locus\_tag="SF1676"

## CDS

/note="Residues 1 to 171 of 171 are 99 pct identical to residues 29 to 199 of a 199 aa protein from Escherichia coli K12 ref: NP\_416166.1"

/codon\_start=1  
/transl\_table=11

## CDS

/product="orf, conserved hypothetical protein"

/protein\_id="AAN43257.1"

## CDS

/db\_xref="GI:24052008"

## CDS

/translation="WGLSELLKTAEVKPGSFYHFRSKEAFGAMLERHYAAYHQRLA  
ELLQSGNRYDRILAYVQOTNQFCOHTISGGLTVKLSAEVCDSLSEDMRSAMDKGA  
RGVIALLSQALENGENHCLTFCGEPLQQAQVLYALWLGANLQAKISRSPFELENALA  
HVKNLIATPAV"

5762. .6859  
/gene="nemA"

## CDS

/locus\_tag="SF1677"

## CDS

5762. .6859  
/gene="nemA"

## CDS

/locus\_tag="SF1677"

## CDS

/note="Residues 1 to 365 of 365 are 99 pct identical to residues 1 to 365 of a 365 aa protein from Escherichia coli K12 ref: NP\_416167.1"

/codon\_start=1  
/transl\_table=11

## CDS

/product="N-ethylmaleimide reductase"

/protein\_id="AAN43258.1"

```
/db_xref="GI:24052009"
/translation="MSSEKLYSLPKVKGALTAANRIFMAPLTRLSIEPGDIPTLMAE
YQRASAGLITSEATQISAOAGYAGAGIHSPEIOIAAWKKITAGVHAENGHMAQL
WHTGRISHASLOPQGOAPVAPSAISAGTSTUSDENGQAIRVETSMPRALEELPIGI
VNDPQAIANAREAPFDLVSLHSAGYLLHQLFSPSSNRHTDOYGSVENARKLVLEV
VDAGLEWAGDRIKIRVSGITPQNTDNGPNEEDADYLIHQKRGKILAYLRMSPDW
AGCEPYTDAPREKVRARFPGIIGAGAYTVEKAETLIGKGLIDAVAFGRDWIANPDLV
ARLQKAEINPQRAESFYGGAGBYDYDPTL"
gene
6940. .7347
/genes="gloA"
/locus_tag="SF1678"
6940. .7347
/genes="gloA"
/locus_tag="SF1678"
/notes="Residues 1 to 135 of 135 are 100 pct identical to
residues 1 to 135 of a 135 aa protein from Escherichia
coli O157:H7 EDL933 ref: NP_288087.1"
/codon_start=1
/transl_table=11
/product="gloA gene product"
/protein_id="AAN43259.1"
/db_xref="GI:24052010"
/translation="WRLHLMRLVGVGLQRSIDFYTKVLGKLLRTSENPEYKSLAFV
GIGPTEEAIVIELTYNWGDKYELGTAYGHIALSVDNAACEKIQNGNVNTRVAGP
VKGTTVIAFVEDPDGKYKLEIEEKDAGRLGN"
gene
7450. .8097
/genes="rnt"
/locus_tag="SF1679"

Alignment Scores:
Pred. No.: 5,41e-26 Length: 14203
Score: 476.50 Matches: 92
Percent Similarity: 68.24% Conservative: 24
Best Local Similarity: 54.12% Mismatches: 51
Query Match: 50.16% Indels: 3
DB: 1 Gaps: 2

US-10-009-916A-1 (1-180) x AE015188 (1-14203)
QY 10 SerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerVal 29
Db 3746 AGTCTGGCCATTCTGGCGCTGGTT-----GTTGGCGCGCGGCACCAAGTCGCATGAA 3693
QY 30 ThrSerGluValHisMetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 3692 AAAGTCAGATGAACCTCGTCACGTCGCAAGGGTTGGGCGTCAATTGCTACGTCACC 3633
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGly 69
Db 3632 ATTACTGAACCATTAAGGCTCTGGAGTTTTCACCGCATCGAAGCATTTACCCCGCGT 3573
QY 70 GluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHis 89
Db 3572 GAACATGGCTTCATATTATGCAAGGAAGCTGCCAGCCACCAAGATGGCAAA 3513
QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLysHisGlu 109
Db 3512 GCCAGCGCGCGGAATCCGCGAGCGGCGCATCTTGATCCACAAAATACCGGTAAACATGAA 3453
QY 110 GlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValLysAlaAspGly 129
Db 3452 GGCCCAAGAGGTACCGGCGCATTTAGCGCATCTCGCTGCATCGTCAATAATGACGGC 3393
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLysGluIleLysGlyArg 148
Db 3392 AAAGTCACCATCGCGTCATCGCGCGCTCGTCAAAATCACTGGATGAATCAAGACAA 3333
QY 149 ThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLysProLeuGlyGly 168
Db 3332 CGCGTATGTCACCGTTGGCGCGCATTAATATCGCATCAACTAAACCGCTGGCGCGT 3273
QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
```

```
Db 3272 GGCGGTGAACGCTATGCCTGTGTGTAATT 3243
RESULT 2
AE016983/c
LOCUS Shigella flexneri 2a str. 2457T section 6 of 16 of the complete
DEFINITION genome.
ACCESSION AE016983
VERSION AE016983.1 GI:30041148
KEYWORDS Shigella flexneri 2a str. 2457T
SOURCE Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
ORGANISM Enterobacteriaceae; Shigella.
REFERENCE 1 (bases 1 to 290628)
AUTHORS Wei,J., Goldberg,M.B., Burland,V., Venkatesan,M.M., Deng,W.,
Fournier,G., Mayhew,G.F., Plunkett,G. III, Rose,D.J., Darling,A.,
Mau,B., Perna,N.T., Payne,S.M., Runyen-Janecky,L.J., Zhou,S.,
Schwartz,D.C. and Blattner,F.R.
TITLE Complete Genome Sequence and Comparative Genomics of Shigella
flexneri Serotype 2a Strain 2457T
JOURNAL Infect. Immun. 71 (5), 2775-2786 (2003)
PUBMED 12704152
REFERENCE 2 (bases 1 to 290628)
AUTHORS Wei,J., Goldberg,M.B., Burland,V., Venkatesan,M.M., Deng,W.,
Fournier,G., Mayhew,G.F., Plunkett,G. III, Rose,D.J., Darling,A.,
Mau,B., Perna,N.T., Payne,S.M., Runyen-Janecky,L.J., Zhou,S.,
Schwartz,D.C. and Blattner,F.R.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2002) Genetics Laboratory, University of
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
1. .290628
/organism="Shigella flexneri 2a str. 2457T"
/mol_type="genomic DNA"
/strain="2457T"
/serotype="2a"
/db_xref="taxon:198215"
93. .1085
/genes="tehA"
/locus_tag="S1493"
93. .1085
/genes="tehA"
/locus_tag="S1493"
/function="transport; Drug/analog sensitivity"
/notes="residues 1 to 330 of 330 are 94.24 pct identical to
residues 1 to 330 of 330 from Escherichia coli K-12 :
B1429"
/codon_start=1
/transl_table=11
/product="Tellurite resistance protein TehA"
/protein_id="AAP16878.1"
/db_xref="GI:30041149"
/translation="MQSDKVLNLDPAGYFVIGTGMGFAPYASQVQVSHWLGDL
VILAMIWGLTSATITRIRFPHSAVEHRPVLSSFVSLFPAITMLVAIGFVWFR
PLACMFSGVTVQAYATAGLWRGSHPEATPGLYLTPANNFTSAMACGALG
YTDAGLVFAGVFWLSLEVLQIRSSGLPTALRTSLGTLQPLAVACSAWLSV
NGSGDTLAKMLFGYLLQLLELMPWLSQPPNAPSPFSRGSVALATTGLHGS
GSDNFFHTLAVPLFIPTFTFIALLIRTFALLMQGKLLVTRERAVLKAEDKE"
1082. .1675
/genes="tehB"
/locus_tag="S1494"
1082. .1675
/genes="tehB"
/locus_tag="S1494"
/function="putative transport; Drug/analog sensitivity"
/notes="residues 1 to 197 of 197 are 98.98 pct identical to
residues 1 to 197 from Escherichia coli K-12 :
B1430"
/codon_start=1
/transl_table=11
/product="Tellurite resistance protein TehB"
/protein_id="AAP16879.1"
```



QY 30 ThrSerGluValHisMetLeuAspAsnGlyIleLysGlnSerIleGlyThrValThr 49  
 Db 287391 AAGTCCAGATGAACCTCGTCACGTCGAAGGGTTGGGCAGTCAATTGGTAGCGTCACC 287332  
 QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGly 69  
 Db 287331 ATTACTGAACCCGATAAAGTCTGGAGTTTACCCCGATCTGAAGCATTTACCCCGCGGT 287272  
 QY 70 GluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHis 89  
 Db 287271 GAACATGGCTTCATATTATGCAAGGAGTGCAGCCAGCCACCAAGATGGCAA 287212  
 QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLysHisGlu 109  
 Db 287211 GCGAGCGCGCGGAATCCGCGAGCGGCATCTTGATCCACAAATACCGGTAAACATGAA 287152  
 QY 110 GlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129  
 Db 287151 GGGCCAGAGGTACCGGCGATTTAGCGCATCTGCCCTGCATCGTCAATAATGACGGC 287092  
 QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLysGluLysGlyArg 148  
 Db 287091 AAGCTACCGATGCCGTCATCGCGCTCGTCTGAATCACTGGATGAATCAAGACAAG 287032  
 QY 149 ThrValMetIleHisAlaGlyAspAsnTyrSerAspLysProLeuProLeuGlyGly 168  
 Db 287031 GCGCTGATGTCACGCTGGCGCGATATATGTCGATCAACTAAACCGCTGGCGGT 286972  
 QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178  
 Db 286971 GCGCGTGAACCGTATCGCTGTGTGAATT 286942  
 RESULT 3  
 ECUS1242 919 bp DNA linear BCT 12-JUL-1996  
 LOCUS Escherichia coli copper-zinc superoxide dismutase (sodC) gene,  
 DEFINITION complete cds.  
 ACCESSION U51242  
 VERSION U51242.1 GI:1256445  
 KEYWORDS  
 SOURCE Escherichia coli  
 ORGANISM Escherichia coli  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.  
 REFERENCE 1 Imlay, K.R. and Imlay, J.A.  
 AUTHORS Imlay, K.R. and Imlay, J.A.  
 TITLE Cloning and analysis of sodC, encoding the copper-zinc superoxide  
 dismutase of Escherichia coli  
 JOURNAL J. Bacteriol. 178 (9), 2564-2571 (1996)  
 MEDLINE 96196162  
 PUBMED 8626323  
 REFERENCE 2 (bases 1 to 919)  
 AUTHORS Imlay, K.R.C. and Imlay, J.A.  
 DIRECT SUBMISSION  
 TITLE Submitted (13-MAR-1996) K.R.C. Imlay, Microbiology, University of  
 Illinois at Urbana-Champaign, 131 Burrill Hall, 407 S. Goodwin  
 Ave., Urbana, IL 61801, USA  
 FEATURES Location/Qualifiers  
 1. .919  
 /organism="Escherichia coli"  
 /mol\_type="genomic DNA"  
 /strain="W3110"  
 /db\_xref="taxon:562"  
 /map="37.1"  
 /note="Subcloned from clone #317 of the Kohara lambda  
 library."  
 224. .745  
 /gene="sodC"  
 224. .745  
 /gene="sodC"  
 /EC\_number="1.15.1.1"  
 /codon\_start=1  
 /transl\_table=11

/product="copper-zinc superoxide dismutase"  
 /protein\_id="AA03729.1"  
 /db\_xref="GI:1256446"  
 /translation="MKRSLAILALVAVATGAQAASEKVEMLNLTSGVQSGISVIT  
 ETDKGLFSPDLKALPNDGHHIHKSCQPATKDKASAAESAGHLDLPQNTGHE  
 GPEGAGHLGDLFALVVDNDGATDAVIAPRLKSLDEIKDKALMVHVGGDNMSDPKPL  
 GGGGERYACGVIK"  
 ORIGIN  
 Alignment Scores: 5.49e-27 Length: 919  
 Pred. No.: 474.50 Matches: 92  
 Score: 68.24% Conservative: 24  
 Percent Similarity: 54.12% Mismatches: 51  
 Best Local Similarity: 49.95% Indels: 3  
 Query Match: 2  
 DB: 1  
 US-10-009-916A-1 (1-180) x ECUS1242 (1-919)  
 QY 10 SerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerVal 29  
 Db 236 AGTCTGGCTATTCTGGCGTGGTT-----GTGCAACCGCGCCACAAGCTGCCAGTGAA 289  
 QY 30 ThrSerGluValHisMetLeuAspAsnGlyIleLysGlnSerIleGlyThrValThr 49  
 Db 290 AAAGTCAGATGAACCTCGTCACGTCGCAAGGGGTAGGCGAGTCAATTGGTAGCGTCACC 349  
 QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGly 69  
 Db 350 ATTACTGAACCCGATAAAGTCTGGAGTTTTCGCCCGATCTGAAAGCATTTACCCCGCGGT 409  
 QY 70 GluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHis 89  
 Db 410 GAACATGGCTTCATATTATGCAAGGAGTGCAGCCAGCCACCAAGATGGCAA 469  
 QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLysHisGlu 109  
 Db 470 GCGAGCGCGCGGAATCCGCGAGCGGCATCTTGATCCACAAATACCGGTAAACATGAA 529  
 QY 110 GlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129  
 Db 530 GGGCCAGAGGTACCGGCGATTTAGCGCATCTGCCCTGCATCGTCAATAATGACGGC 589  
 QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLysGluLysGlyArg 148  
 Db 590 AAGCTACCGATGCCGTCATCGCGCTCGTCTGAATCACTGGATGAATCAAGACAAA 649  
 QY 149 ThrValMetIleHisAlaGlyAspAsnTyrSerAspLysProLeuProLeuGlyGly 168  
 Db 650 GCGCTGATGTCACGCTGGCGCGATATATGTCGATCAACCTAAACCGCTGGCGGT 709  
 QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178  
 Db 710 GCGCGTGAACCGTATCGCTGTGTGAATT 739  
 RESULT 4  
 AE005387/c 12347 bp DNA linear BCT 21-MAR-2001  
 LOCUS Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 6 of  
 DEFINITION 290.  
 ACCESSION AE005387  
 VERSION AE005387.1 GI:12515625  
 KEYWORDS  
 SOURCE Escherichia coli O157:H7 EDL933  
 ORGANISM Escherichia coli O157:H7 EDL933  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.  
 REFERENCE 1 (bases 1 to 12347)  
 AUTHORS Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D.,  
 Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A.,  
 Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L.,  
 Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamousis, K.,  
 Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C.,

```

Welch,R.A. and Blattner,F.R.
Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
Nature 409 (6819), 529-533 (2001)
21074935
11206551
2 (bases 1 to 12347)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grotbeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potancusis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Direct Submission
Submitted (22-OCT-2000) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
Location/Qualifiers
1..12347
/organism="Escherichia coli O157:H7 EDL933"
/mol_type="genomic DNA"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:155864"
/note="enterohaemorrhagic"
336..1838
/gene="ydgR"
/note="synonym: Z2646"
336..1838
/gene="ydgR"
/function="putative transport; Not classified"
/note="Residues 1 to 500 of 500 are 99.80 pct identical to
residues 1 to 500 of 500 from Escherichia coli K-12 Strain
MG1655: B1634"
MG1655: B1634"
/codon_start=1
/transl_table=1
/product="putative transport protein"
/db_xref="GI:12515626"
/gene="gst"
1944..2549
/note="synonym: Z2647"
1944..2549
/gene="gst"
/function="enzyme; Biosynthesis of cofactors, carriers:
Thioredoxin, glutaredoxin, glutathione"
/note="Residues 1 to 201 of 201 are 100.00 pct identical
to residues 1 to 201 of 201 from Escherichia coli K-12
Strain MG1655: B1635"
/codon_start=1
/transl_table=1
/product="glutathione S-transferase"
/protein_id="AAG56624.1"
/db_xref="GI:12515627"
/translacion="MKLPYKPGACSLASHITLRESGKDTFLVSVLDMKKRLKLENGDDYF
AVNPKGVFALLDDTLTGVAIMOVLADSPVDQLLAPVNSISRYKTIEWLNLYIA
TELHKGTPFPPEYKPTVRAQLEKLYVNEALKDEHWICGQRTIADAYLFT
VLRWAYAVKLNLEGLHIAEFMORMAERPEVQDLSAEGLK"
complement(2593..3456)
/gene="pdxY"
/note="synonym: Z2648"
complement(2593..3456)
/gene="pdxY"
/function="enzyme; Biosynthesis of cofactors, carriers:
Pyridoxine"
/note="Residues 1 to 287 of 287 are 99.65 pct identical to

```

```

residues 1 to 287 of 287 from Escherichia coli K-12 Strain
MG1655: B1636"
/codon_start=1
/transl_table=11
/product="pyridoxal kinase 2 / pyridoxine kinase"
/protein_id="AAG56625.1"
/db_xref="GI:12515628"
/translacion="MMKNILAIQSHVVYGHAGNSAAEPMPRLGANVPLNTVQFSNH
TOYCKWTCVMPSPSHLTELIOGIAIDKLHCTDAVLSCYLGSAEGGHEILGIVROVKA
NOYKQVPCDPMVGHGPKGCIIVAGVAEPVHRHGLPASDIAPNLVLEILEICHEPVNN
VEEAVLAARELTAGQPOIVLVKHLARAGYSRDRFEMLLVLTADAEAWHISRPLVDFGNRQ
PVGVDTVTSGLLVKLLQGATLEALHEHTAAVYIEMVTIKAMQEIYELQVVAQAQRIA
KPEHYFSATKL"
complement(3515..4801)
/gene="tyrS"
/note="synonym: Z2650"
complement(3515..4801)
/gene="tyrS"
/function="enzyme; Amino acyl tRNA syn; tRNA modifc'n"
/note="Residues 5 to 428 of 428 are 100.00 pct identical
to residues 1 to 424 of 424 from Escherichia coli K-12
Strain MG1655: B1637"
/codon_start=1
/transl_table=11
/product="tyrosine tRNA synthetase"
/protein_id="AAG56626.1"
/db_xref="GI:12515629"
/translacion="MEILMASSNLIKQLOERGLVAQVTDERALAERLAQGPALYCGF
DPTADSLHLGVPLLLCKRQOAGHPVALVGSATGLIGDPSKAAERKLNTEETVQ
EWDKIRKQVAPFFDFDCGNSAIAANNYDFWNNVLTFLURDIGKHSVQNMINKEA
VKNQNRDQGISFTFESYNLLQGYDFACLNKQYGVLLIGGSDQMGWNTSGIDLTRR
LHQNVFGLTVPLTIKADGTGFKTEGAVWLDPKTSPYFYQFWINTADADVYRFL
KFTFMSIEINALDEEDKNSGAPRAQYVLAQVTRLVHGEGLQAARITECLFSG
SLSALSEADPEQLAODGVPMVMEKGLDMLQALVDELQPSRQARKTIASWAITING
EKQSDPYFFKEEDRLGRFTLLRRGKKNYCLCWK"
complement(4919..5575)
/gene="pdxH"
/note="synonym: Z2652"
complement(4919..5575)
/gene="pdxH"
/function="enzyme; Biosynthesis of cofactors, carriers:
Pyridoxine"
/note="Residues 1 to 218 of 218 are 99.54 pct identical to
residues 1 to 218 of 218 from Escherichia coli K-12 Strain
MG1655: B1638"
/codon_start=1
/transl_table=11
/product="pyridoxinephosphate oxidase"
/protein_id="AAG56627.1"
/db_xref="GI:12515630"
/translacion="MSDNDELQQAHLRREYTKGLRRRDLPADPLTLFERWLSQACE
AKLADPTAMVATVDEHGQPYORIVLLKHVDEKGMVFYTNLGSRAKAOIENNPVYSL
FPWHTLREQVMVICKAERLSTLEVMKYFHSRPSDQIGSWYSKQSSRISARGILBSKF
LELKQKQQGEVPLPFSFWGFRVSLQIEFWQGGHEHLRDLFQRENDAWKIDLAP"
complement(5634..5882)
/gene="ydhA"
/note="synonym: Z2653"
complement(5634..5882)
/gene="ydhA"
/function="orf; Unknown function"
/note="Residues 1 to 82 of 82 are 96.34 pct identical to
residues 1 to 82 of 82 from Escherichia coli K-12 Strain
MG1655: B1639"
/codon_start=1
/transl_table=11
/protein_id="AAG56628.1"
/db_xref="GI:12515631"
/translacion="MOTDTLEYQDEKPLTVKLNPNRQSVFVYDNLHLKQGLSAS
GARYSDGIYVFWFSKGEATVYKRDRIVLNNQLQNPQR"
complement(6061..7170)
/gene="Z2654"
complement(6061..7170)

```

```
/gene="Z2654"
/function="orf; Unknown function"
/note="Residues 1 to 368 of 369 are 99.45 pct identical to
residues 1 to 368 of 369 from Escherichia coli K-12 Strain
MG1655: B1640"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG56629.1"
/db_xref="GI:12515632"
/translation="MKSGRFIGVMSGTSLDGVVDVVLATIDHRVLAQLSLWPIPVSL
KQAVDLCQQLTQLSQFOLDTQUGRLFADAVKALKQBNLOARDIVAIGHGQTVW
HEPTGVAPHTLQIGDNNQIVARTGITVVDGFRRRDIALGQGGAPLVPFAHLLAHP
EERVLNIGLILIPQVPVGGYDTGPNMLMDAMTRQAGKPYDKDAEWARAGK
VILPLLQNLSPDYPISOPAKPSGTREYFNKWLRLHRHPGVDPRDVAQLAELTAV
TISEVILSGGGERLWVCGGSRNPLLMARLAALLPGTEVTITDAVGISDDMEALAF
AMLAWRTLAGLPGLPSVIGASQETVLGALFFANS"
7444..7911
/gene="slyB"
/note="synonym: Z2655"
7444..7911
/gene="slyB"
/function="putative membrane; Not classified"
/note="Residues 1 to 155 of 155 are 99.35 pct identical to
residues 1 to 155 of 155 from Escherichia coli K-12 Strain
MG1655: B1641"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/protein_id="AAG56630.1"
/db_xref="GI:12515633"
/translation="MKRVLVSVMSVGLSVGVNNDLSDGVYVTAEEKQVQNVSYGT
IVNRPVQIQGGDDSNVIGAVGLGGFLGNTVGGTGRSLATAGAVAGVAGGV
QSAMNKTQGVLEIRKDDQNTIMVQKQNTREFSPQRVVLASNGSQVTVSPR"
complement(7958..8398)
/gene="slyA"
/note="synonym: Z2657"
complement(7958..8398)
```

## Alignment Scores:

```
Pred. No.: 6 69e-26 Length: 12347
Score: 474.50 Matches: 92
Percent Similarity: 68.24% Conservatives: 24
Best Local Similarity: 54.12% Mismatches: 51
Query Match: 49.95% Indels: 3
DB: Gaps: 2
```

US-10-009-916a-1 (1-180) x AE005387 (1-12347)

```
QY 10 SerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerVal 29
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 12211 AGTCTGGCATTCTGGCGTGGTT-----GTTGCGACGGCGGCACAGCTGCACGTGAA 12158

QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 12157 AAGATCGAGATGAACCTGCACGTGCGCAAGGGTTGGGCAGTCACATTCGTGCGTCAAC 12098

QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGly 69
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 12097 ATTACTGAACCATGAAGTCTGGAGTTTTCGCCGATCTGAAGCATTTACCTCCCGGT 12038

QY 70 GluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHis 89
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 12037 GAACATGGCTTCATTATTCATGCAAGGAAGCTGCCAGCAGCCACCAAGATGGCAA 11978

QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisTyAspProAspLysThrGlyLysHisGlu 109
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 11977 GCCAGCGCGCGGAATCCGAGCGCGGCATCTTGTATCCAAAAATACCGTAAACATGAA 11918

QY 110 GlyProLeuGlyAsnGlyLysGlyAspLeuProArgLeuValValLysAlaAspGly 129
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 11917 GGGCCAGAAAGTCCCGGCAATTTAGGCGATCTCGCTGCATCGTCTCAATATGACGCG 11858
```

```
QY 130 IleAlaLysGluThrIleuLeuAlaProArgLeu---ThrValLysGluIleLysGlyArg 148
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 11857 AAGACTACCGATCGCGTCACTCGCTCTGAAATCACTGGATCAATCAAGACAAA 11798

QY 149 ThrValMetIleHisAlaGlyGlyAspAsnTySerAspLysProLeuProLeuGlyGly 168
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 11797 GCGCTGATGTCACGTTGGCGCGCATATAATATGTCGATCAACCTAAACCGCTGGCGGT 11738

QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 11737 GCGCGTGAACGCTATGCGCTGTGTGTATTT 11708
```

## RESULT 5

U00096 17/c

WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| U00096_00     | 1       | 110000  |
| U00096_01     | 100001  | 210000  |
| U00096_02     | 200001  | 310000  |
| U00096_03     | 300001  | 410000  |
| U00096_04     | 400001  | 510000  |
| U00096_05     | 500001  | 610000  |
| U00096_06     | 600001  | 710000  |
| U00096_07     | 700001  | 810000  |
| U00096_08     | 800001  | 910000  |
| U00096_09     | 900001  | 1010000 |
| U00096_10     | 1000001 | 1110000 |
| U00096_11     | 1100001 | 1210000 |
| U00096_12     | 1200001 | 1310000 |
| U00096_13     | 1300001 | 1410000 |
| U00096_14     | 1400001 | 1510000 |
| U00096_15     | 1500001 | 1610000 |
| U00096_16     | 1600001 | 1710000 |
| U00096_17     | 1700001 | 1810000 |
| U00096_18     | 1800001 | 1910000 |
| U00096_19     | 1900001 | 2010000 |
| U00096_20     | 2000001 | 2110000 |
| U00096_21     | 2100001 | 2210000 |
| U00096_22     | 2200001 | 2310000 |
| U00096_23     | 2300001 | 2410000 |
| U00096_24     | 2400001 | 2510000 |
| U00096_25     | 2500001 | 2610000 |
| U00096_26     | 2600001 | 2710000 |
| U00096_27     | 2700001 | 2810000 |
| U00096_28     | 2800001 | 2910000 |
| U00096_29     | 2900001 | 3010000 |
| U00096_30     | 3000001 | 3110000 |
| U00096_31     | 3100001 | 3210000 |
| U00096_32     | 3200001 | 3310000 |
| U00096_33     | 3300001 | 3410000 |
| U00096_34     | 3400001 | 3510000 |
| U00096_35     | 3500001 | 3610000 |
| U00096_36     | 3600001 | 3710000 |
| U00096_37     | 3700001 | 3810000 |
| U00096_38     | 3800001 | 3910000 |
| U00096_39     | 3900001 | 4010000 |
| U00096_40     | 4000001 | 4110000 |
| U00096_41     | 4100001 | 4210000 |
| U00096_42     | 4200001 | 4310000 |
| U00096_43     | 4300001 | 4410000 |
| U00096_44     | 4400001 | 4510000 |
| U00096_45     | 4500001 | 4610000 |
| U00096_46     | 4600001 | 4639675 |

Continuation (18 of 47) of U00096 from base 1700001 (U00096 Escherichia coli K-12 MG1655)

| Alignment Scores:      | 5.48e-25 | Length:        | 110000 |
|------------------------|----------|----------------|--------|
| Pred. No.:             | 474.50   | Matches:       | 92     |
| Score:                 | 68.24%   | Conservatives: | 24     |
| Percent Similarity:    | 54.12%   | Mismatches:    | 51     |
| Best Local Similarity: | 49.95%   | Indels:        | 3      |
| Query Match:           | 1        | Gaps:          | 2      |



```
US-10-009-916A-1 (1-180) x U00096_17 (1-110000)

Qy 10 SerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerVal 29
Db 22667 AGTCTGGCTATTCGGCGCTGGT-----GTTGCAACCGCGGCACAGCTGCCAGTGAA 22614
Qy 30 ThrSerGluValHisMetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 22613 AAAGTCGAGATCAACTCGTCACGTCGCAAGGGGTAGGCAGTCAATTGGTAGCGTCACC 22554
Qy 50 PheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGly 69
Db 22553 ATTACTGAACCGATAAAGGCTCGAGTTTTCGCCCGATCTCAAGAGCATTACCCCGGT 22494
Qy 70 GluHisGlyPheHisIleHisGluGlySerCysGlyProAlaGluHisAspGlyHis 89
Db 22493 GAACATGCTCTCCATATTATTCGCCAAAGAAAGCTCCAGCCAGCCACCAAGATGGCAA 22434
Qy 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisIleValAspProAspLysThrGlyLysHisGlu 109
Db 22433 GCCAGCGCGCGGAATCCGCGGGGCGATCTTGATCCACAAATACCGGTAAACATGAA 22374
Qy 110 GlyProLeuGlyAsnGlyHisGlyAspLeuProArgLeuValValIleAlaAspGly 129
Db 22373 GGGCCAGAAGGTGCGCGGCATTAGGCGATCTGCCCTGCACCTGCTCGTCAATAATACGCGC 22314
Qy 130 IleAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLysGluIleLysGlyArg 148
Db 22313 AAAGCTACCGATCGCGCTCGCGCGCTCGTCTGAAATCACTGATGAATCAAGACAA 22254
Qy 149 ThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGly 168
Db 22253 GCGCTGATGTCACCGCTTGGCGCGCATATATGTCGATCAACCTAAACCGCTGGCGGT 22194
Qy 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
Db 22193 GCGCGTGAACGCTATGCTGTGTGTAATT 22164

RESULT 6
AP002558/c AP002558 257071 bp DNA linear BCT 20-MAR-2004
LOCUS Escherichia coli O157:H7 DNA, complete genome, section 9/20.
DEFINITION AP002558 BA000007
ACCESSION AP002558.1 GI:13361764
VERSION
KEYWORDS
SOURCE Escherichia coli O157:H7
ORGANISM Escherichia coli O157:H7
REFERENCE 1 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
AUTHORS 1 Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T., Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasaki, C., and Shinagawa, H.
TITLE Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak
GENES Genet. Syst. 74 (5), 227-239 (1999)
JOURNAL 20198780
MEDLINE 10734605
PUBMED
REFERENCE 2
AUTHORS 2 Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M., Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and Hayashi, T.
TITLE Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an Escherichia coli K-12 strain MG1655
JOURNAL Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
MEDLINE 20557356
PUBMED 11108008
REFERENCE 3
AUTHORS 3 Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S.,
```

```
Yutsudo, C.H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T., Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasaki, C. and Shinagawa, H.
Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak
Gene 258 (1-2), 127-139 (2000)
MEDLINE 20564182
PUBMED 11111050
AUTHORS 4 Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasaki, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and Shinagawa, H.
Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12
DNA Res. 8 (1), 11-22 (2001)
MEDLINE 21156231
PUBMED 11258796
AUTHORS 5 (bases 1 to 257071) Hattori, M., Ishii, K. and Shiba, T.
Direct Submission
Submitted (26-JUN-2000) Masahira Hattori, Kitasato Institute for Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan
(E-mail: hattori@genome.its.kitasato-u.ac.jp, URL: http://genome.its.kitasato-u.ac.jp/, Tel: 81-42-778-8194, Fax: 81-42-778-8193)
genome project
This work was done in collaboration with Tetsuya Hayashi, Makoto Ohnishi, Keisuke Nakayama (Miyazaki Medical College), Kozo Makino, Ken Kurokawa, Katsumi Yokoyama, Masashi Tanaka, Takeshi Honda, Teruo Yasunaga, Hideo Shinagawa (Osaka University), Takahiro Murata (Shinsu University), Chang-Gyun Han, Eiichi Ohtsubo, Toru Tobe, Chihiro Sasaki (University of Tokyo), Hideto Takami (Japan Marine Science and Technology Center), Naotake Ogasawara (Nara Institute of Science and Technology), Satoru Kuhara (Kyuushu University), and supported by the Research for the Future Program of the Japan Society for the Promotion of Science.
FEATURES
Location/Qualifiers
1..257071
/organism="Escherichia coli O157:H7"
/mol_type="genomic DNA"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:83334"
93..1409
/gene="ECs2298"
93..1409
/gene="ECs2298"
/note="similar to B1592_ECOLI gi|1787876 percent identity 99 in 438 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative chloride channel"
/protein_id="BAB35721.1"
/db_xref="GI:13361765"
/translation="MFHCFLLMHLLHTYPLRTMFRLLIATVVGILAAFAVAGFRH
AMLLELLEFLNNDGSLVNAATNISPMRLLTTPALGGLAAGLLMGKQFTQORPHAP
RQEWKLTACGAAAGAAARAPLAGSLFAEVLFGTMMLASLGFVTSVAVLLVSN
LQNHSDALLVQLSVTVQARDYALITSTVLGICGLPILLTLMNACHRGVSLKASG
PWLALGLIVGLSLFTPAVMGNGYSTVQSFLTAPPLMIAGFLCKLCVLAASSG
SGACGGVTFPLFIGLAIGLMYGRSLGIMFPDGEIITLLGLTGATLLATTHAPIM
STLMCEMTGEYQLPGLLIACVAVISRTIHRDSDIYRQTHQS"
complement(1362..2069)
/gene="ECs2299"
complement(1362..2069)
/gene="ECs2299"
/note="similar to B1593_ECOLI gi|1787877 percent identity 99 in 235 aa (Conserved in E.coli K-12)"
```





|   |  |  |   |
|---|--|--|---|
| ACCESSION   | AE016761 AE014075  |  | /product="Hypothetical oxidoreductase ydfI"   |
| VERSION   | AE016761.1 GI:26108223   |  | /protein_id="AA080428.1"  |
| KEYWORDS  |  |  | /db_xref="GI:26108228"  |
| SOURCE  | Escherichia coli CFT073  |  | /translation="MGNHLISAKATLPVYDRNNLAPRIVHLGFGAFHRAHOGVYADIL                              |
| ORGANISM  | Escherichia coli CFT073  |  | ATEHSPDWGYEYNLIAGEQIADLQQDNLYTVAEMSADATATATVAVGVVKALHVOI                                |
| Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; |  |  | DGLSTVIAAMCEPOLAIVSLTITKGFHSPATQQLMDHDPVADVNQHPQKPTATG                                  |
| Enterobacteriaceae; Escherichia.                                  |  |  | VIVBALARRKAAGLPAFTVMSCDNMPENGVHVRDVTSYAQIDVLAQWIEDNVTFP                                 |
| REFERENCE   | 1 (bases 1 to 300475)  |  | STVDRITVPAVTEDTLAKIEQITGRDAAGVACEFFQWVEDNFVAGRPPEWAKAGAE                                |
| AUTHORS   | Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P., Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J., Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C., Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R. |  | LVSVDLPEEMKRLMGLSHFLAYGLAGYOHINDCMEDEHVRHAAVTLMLOBPAP                                   |
| TITLE   | Extensive Mosaic Structure Revealed by the Complete Genome Sequence of Uropathogenic Escherichia coli  |  | TLKQGVLDQDYANRLIERYNSPALRHTWQIAMDQSKLPQRMLDVSRWHLAHDSE                                  |
| JOURNAL   | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)  |  | DLLALGVAGNMRYVGGVDEQGNPIEISDPLLPVLOKAVQSSAGTARVQSLAIKALF                                |
| PUBMED  | 12471157   |  | GGDIPGNSLFTTKVTEAYLSLAHGAKATVAKISVK   |
| REFERENCE   | 2 (bases 1 to 300475)  |  | /gene="ydfU"  |
| AUTHORS   | Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P., Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J., Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C., Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R. |  | /locus_tag="c1969"  |
| TITLE   | Direct Submission  |  | complement (2726. .4093)  |
| JOURNAL   | Submitted (20-JUN-2002) Genetics Laboratory, University of Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA   |  | /gene="ydfU"  |
| FEATURES  | Location/Qualifiers  |  | complement (2726. .4093)  |
| source  | 1. 300475  |  | /locus_tag="c1969"  |
| gene  | /organism="Escherichia coli CFT073"  |  | /function="putative transport"  |
| CDS   | /mol_type="genomic DNA"  |  | /note="Escherichia coli K-12 ortholog: b1543; Escherichia coli O157:H7 ortholog: z2153" |
|   | /strain="CFT073"   |  | /codon_start=1  |
|   | /db_xref="taxon:199310"  |  | /transl_table=11  |
|   | 75. .761   |  | /product="Hypothetical metabolite transport protein ydfU"                               |
|   | /gene="ydfH"   |  | /protein_id="AA080429.1"  |
|   | /locus_tag="c1966"   |  | /db_xref="GI:26108227"  |
|   | 75. .761   |  | /translation="MTIEKHERSTKDLVAKAAVSGWGLTALBFMDPQLYSLGAALVFHE                             |
|   | /gene="ydfH"   |  | IFPESSTAMALILAMGTYGAGYVARIVGAFIFGKMGDRIGRKKVLFITITMGICTT                                |
|   | /locus_tag="c1966"   |  | LIGVLPYAOIGVFAPILLYTLRIQLGAGAEISGAGTMAEYAPKGRGILISFVA                                   |
|   | /note="Escherichia coli K-12 ortholog: b1540; Escherichia coli O157:H7 ortholog: z2157"  |  | MGTNCGHLSATAIWAFFMFFILSKBELLAHWRIFPLASVYVYVFWALWRLMKESPVF                               |
|   | /codon_start=1   |  | EKVNDSNPTAKPAGMFSQSFATLGRFGOAGNSGLIQTLFAGLYVQTLDFNK                                     |
|   | /transl_table=11   |  | AIPTDALMISSILGEMTIPFLGMSDKIGRRIPYIIMNTSAIVLAWPLMSIIVDKSYA                               |
|   | /product="Hypothetical transcriptional regulator ydfH"   |  | PSITWVALIVHNCALVGLFALENTMAEMFGCKRFRMAISKEIGLIIASGFGPIL                                  |
|   | /db_xref="GI:26108224"   |  | AGICTWTSWYPIALIMAYSVIGLISALKMPEVKORDLSALEDAEDQPHVVRAAQ                                  |
|   | SVNRSQVLEONLHQRIALERKQDDFFELDDNFHQLITQADCOLANDTLENLK   |  | BSRSL"  |
|   | ATVDRVRYSPDHVSPPPEMLRLQHLDIQSALQKRDGDAVAMTQHLQISESVRQIR  |  | complement (4151. .5194)  |
|   | QENSDFSEEN"  |  | /gene="rspB"  |
|   | 938. .1141   |  | /locus_tag="c1970"  |
|   | /gene="ydfZ"   |  | complement (4151. .5194)  |
|   | /locus_tag="c1967"   |  | /gene="rspB"  |
|   | 938. .1141   |  | /locus_tag="c1970"  |
|   | /gene="ydfZ"   |  | /EC_number="1.1.1.1"  |
|   | /locus_tag="c1967"   |  | /function="enzyme; Global regulatory functions"   |
|   | /note="Escherichia coli K-12 ortholog: b1541; Escherichia coli O157:H7 ortholog: z2156"  |  | /note="Escherichia coli K-12 ortholog: b1580"   |
|   | /codon_start=1   |  | /codon_start=1  |
|   | /transl_table=11   |  | /transl_table=11  |
|   | /product="Hypothetical protein ydfZ"   |  | /product="Starvation sensing protein rspB"  |
|   | /protein_id="AA080427.1"   |  | /protein_id="AA080430.1"  |
|   | /db_xref="GI:26108225"   |  | /db_xref="GI:26108228"  |
|   | XTVVVEGCEELAPLIRLGNM"  |  | /translation="MELVRKVMKSIILIEKPNQLSIIEREIPTPSAGVRVVKLAGI                                |
|   | complement (1177. .2637)   |  | CGSDSHIYRGNPFKYPVIGHEFFGVIVDAGEVESARVGERVADVPVSCGHCYP                                   |
|   | /locus_tag="c1968"   |  | CSIGKPNVCTTAVLGVHADGGSFYAVVPKNAWKIPEAVADQVAVMIEPFTTAANY                                 |
|   | /complement (1177. .2637)  |  | TGHQPTENDTVLVGAGPIGLTIVQVLKGVYKNVIVADRIDERLEKAKESGADWA                                  |
|   | /locus_tag="c1968"   |  | INNSQPLGESFAEKGIKPTLIIDACHPSILKEAVTLASPAARIVLMGFSESPSEVI                                |
|   | complement (1177. .2637)   |  | QQGIIGKELSIFFSRNLNANKFPVVVDWLMSKGLIKEKLIITTFDFQHVADALSLELD                              |
|   | /locus_tag="c1968"   |  | QKHCKVLLTFSE"   |
|   | /EC_number="1.1.1.1"   |  | complement (5182. .6429)  |
|   | /function="putative"   |  | /gene="rspA"  |
|   | /note="Escherichia coli K-12 ortholog: b1542; Escherichia coli O157:H7 ortholog: z2155"  |  | /locus_tag="c1971"  |
|   | /codon_start=1   |  | complement (5182. .6429)  |
|   | /transl_table=11   |  | /gene="rspA"  |
|   | /product="Starvation sensing protein rspA"   |  | /locus_tag="c1971"  |
|   | /protein_id="AA080431.1"   |  | /function="phenotype; Global regulatory functions"                                      |
|   | /db_xref="GI:26108229"   |  | /note="Escherichia coli K-12 ortholog: b1581"   |
|   | ATLNGRELVSAYLQDLICPOLIGRDAHRIEDIQWFFYKGYWRGPGVTMGAISAVDM   |  | /codon_start=1  |
|   | ALWDIKAKAANPLYQLLGGASREGVMVYCHTTHGSIDLEADDVAKHOBGLPKAIRVQ  |  | /transl_table=11  |
|   | CGIPGMKITYGMSKGLAYEPATKQWMPESQLWSTKYLDFMPLKDFAVRNKFGFDE  |  | /product="Starvation sensing protein rspA"  |





```

QY      165 ProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db      669 CCGCTGGCGCGCGGACCGCTTACGCTGGCGGTCAAT 710

RESULT 10
AF056931      1274 bp DNA linear BCT 25-JUN-1999
LOCUS      Salmonella typhimurium copper-zinc superoxide dismutase (sodC-2)
DEFINITION      gene, complete cds.
ACCESSION      AF056931
VERSION      AF056931.1 GI:3046979
KEYWORDS      Salmonella typhimurium
SOURCE      Salmonella typhimurium
ORGANISM      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
              Enterobacteriaceae; Salmonella.
REFERENCE      1 (bases 1 to 1274)
AUTHORS      Fang,F.C., DeGroote,M.A., Foster,J.W., Baumlner,A.J., Ochshner,U.,
              Testerman,T., Bearson,S., Giard,J.C., Xu,Y., Campbell,G. and
              Laessig,T.
TITLE      Virulent Salmonella typhimurium has two periplasmic Cu,
              Zn-superoxide dismutases
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 96 (13), 7502-7507 (1999)
MEDLINE      99307439
PUBMED      10377444
REFERENCE      2 (bases 1 to 1274)
AUTHORS      De Groote,M.A., Ochshner,U.A., Xu,Y., Laessig,T.A., Campbell,G.,
              Foster,J.W. and Fang,F.C.
TITLE      Direct Submission
JOURNAL      Submitted (31-MAR-1998) Medicine, University of Colorado Health
              Sciences Center, 4200 E. 9th Ave., Denver, CO 80262, USA
FEATURES      Location/Qualifiers
              1..1274
                 /organism="Salmonella typhimurium"
                 /mol_type="genomic DNA"
                 /strain="14028s"
                 /db_xref="taxon:602"
                 433..954
                 /gene="sodC-2"
                 433..954
                 /gene="sodC-2"
                 /note="similar to Escherichia coli copper-zinc superoxide
                 dismutase encoded by the sequence presented in GenBank
                 Accession Number U51242"
                 /codon_start=1
                 /transl_table=11
                 /product="copper-zinc superoxide dismutase"
                 /protein_id="AAC13559.1"
                 /db_xref="GI:3046980"
                 /translation="MKRLSLAMVTLACAGAAASEKVRNMLVTAQVQSGSIGTVVID
                 ETEGGLKFTPHLKLPPGHEGPHIHANGSCOPAIKDKAVAAEAAGHLDLPNTGKHE
                 GPEGQGLGLPVLVNNDDIASEPVTAPRLKSLDEVDKALMIHVGGDNMSDQPKPL
                 GGGMRVACGVIK"
ORIGIN
Alignment Scores:
Pred. No.:      4,42e-25      Length:      1274
Score:      451.00      Matches:      91
Percent Similarity:      64.94%      Conservative:      22
Best Local Similarity:      52.30%      Mismatches:      49
Query Match:      47.47%      Indels:      12
DB:      1      Gaps:      3

US-10-009-916A-1 (1-180) x AF056931 (1-1274)

QY      12 ValThrIleSerLeuThrSerIleThrSerValValIleAlaCys----- 27
Db      442 TTAAGTTTACGATGCTGACG-----CTGTGGCTGTGGCGGTGGCGAG 486

QY      28 -----SerValThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIle 45
Db      487 GCGCCGACGAGAAATAGATGAATCTGGTACGCGCGCAAGCGGTAGGCGAGTCTATC 546

```

```

QY      46 GlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGly 65
Db      547 GGCACCGTCGTCAATGAAACCGAGCGGCTTAAATTTACCCACACCTTAAAGCG 606

QY      66 LeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGlu 85
Db      607 TTGCCCGCGCGGAGCATGCTTTTCATTCATGCGCAACGGTAGCTGCCAGCCGCGATT 666

QY      86 HisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrsAspProAspLysThr 105
Db      667 AAAGACGGCAAAAGCGGTTGCCGAGAACCGGCTGGTGGTCTCTGGACCCACAAATACC 726

QY      106 GlyLysHisGluGlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValVal 125
Db      727 GGCACATGAGAGCGGAGCGGAGCCAGGCGCATCTGGCGACCTCCGGGTGTAGTCGT 786

QY      126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLysGlu 144
Db      787 AATAATGATGTTATCGCCAGCAACCGGTTACTGCGCGCGTCTGAAGTCTCTTGTATGAA 846

QY      145 IleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrsSerAspLysProLeu 164
Db      847 GTAAAGATAAAGCGCTCATGATCCATGTGGCGCGGATAACATGTCGATCGACCGGAAA 906

QY      165 ProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db      907 CCGCTCGGCGCGCGGAAATGCTTACGCTGGCGGAGTCAAT 948

RESULT 11
AE008762      21913 bp DNA linear BCT 23-APR-2003
LOCUS      Salmonella typhimurium LT2, section 66 of 220 of the complete
              genome.
ACCESSION      AE008762 AE006468
VERSION      AE008762.1 GI:16419944
KEYWORDS
SOURCE
ORGANISM      Salmonella typhimurium LT2
              Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
              Enterobacteriaceae; Salmonella.
REFERENCE      1 (bases 1 to 21913)
AUTHORS      McCrelland,M., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F.,
              Hou,S., Layman,D., Leonard,S., Nguyen,C., Scott,K., Holmes,A.,
              Grewal,N., Mulvaney,E., Ryan,E., Sun,H., Florea,L., Miller,W.,
              Stoneking,T., Nhan,M., Waterston,R. and Wilson,R.K.
              Complete genome sequence of Salmonella enterica serovar Typhimurium
              LT2
              Nature 413 (6858), 852-856 (2001)
JOURNAL      21534948
MEDLINE      11677609
PUBMED
REFERENCE      2 (bases 1 to 21913)
AUTHORS
CONSTRM
TITLE
JOURNAL
COMMENT
              The Salmonella typhimurium Genome Sequencing Project
              Direct Submission
              Submitted (29-MAR-2001) Genome Sequencing Center, Department of
              Genetics, Washington University School of Medicine, 4444 Forest
              Park Boulevard, St. Louis, MO 63108, USA
              COMMENT      Supported by NIH grant 5U 01 AI43283
              Coding sequences below are predicted from manually evaluated
              computer analysis, using similarity information and the programs;
              GLIMMER; http://www.tigr.org/softlab/glimmer/glimmer.html and
              GeneMark; http://opal.biology.gatech.edu/GeneMark/
              EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
              Encyclopedia of Genes and Genomes; http://www.genome.ad.jp/kegg/,
              and Pedro Romero and Peter Karp at EcoCyc;
              http://ecocyc.PangeaSystems.com/ecocyc/
              The analyses of ribosome binding sites and promoter binding sites
              were kindly provided by Heladia Salgado, Julio Collado-Vides and

```

RegulonDB;  
[http://kinich.cifn.unam.mx:8850/db/regulondb\\_intro.frameset](http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset)

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

FEATURES  
 source  
 1..21913  
 /organism="Salmonella typhimurium LT2"  
 /mol\_type="genomic DNA"  
 /strain="LT2; SGSC 1412; ATCC 700720"  
 /db\_xref="ATCC:700720"  
 /db\_xref="taxon:99287"  
 /note="LT2"  
 complement(106..179)  
 /gene="valW"  
 /note="synonym: STM1423"  
 complement(106..179)  
 /gene="valW"  
 /product="tRNA-Val"  
 complement(193..266)  
 /gene="valV"  
 /note="synonym: STM1424"  
 complement(193..266)  
 /gene="valV"  
 /product="tRNA-Val"  
 complement(428..1812)  
 /gene="ydhE"  
 /note="synonym: STM1425"  
 complement(428..1801)  
 /gene="ydhE"  
 /note="similar to E. coli putative transport protein (AAC74735.1); Blastp hit to AAC74735.1 (457 aa), 92% identity in aa 1 - 457"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative MATE family transport protein"  
 /protein\_id="AAL20347.1"  
 /db\_xref="GI:16419945"  
 /translation="MQKTSARQLLARIPIVLAQAOTAMGFVDVTVMAGYSGTDM AAVATGTTWIPAILFGHLLIALTPTVIAQLNGSRRERIAHVRQGFWLAGFVSIVL MVLNAGVITIRSMNIDPALADKAVGLRALLGAPQYLFQVARNOCGLAKTKPG MVMGLGLNIPVNYIETXGHFGMPELGGICGVATAVYWMFIMLSYIKHARSM RDIREKGFGRDPSVMKRLQLGLPIALALFFETVLFPAVALLVSPGLIVDVAGHQI ALNPSLMFVLPMSLAAAVTIKVGRLGGSTLDAQTARTGLGVGICMAVVTALFTV TLKHIALLYNDNPVLAALMLAAVYQISDSIQVIGSILRGYKDTRSIPFTF TAYWVGLPSGYIALTDLVVDVDRMGPFAGFWGFTIGLTSAAVLMMLRMRYLQROPFAI ILQRAAR"  
 complement(1807..1812)  
 /gene="ydhE"  
 /note="putative RBS for ydhE; RegulonDB: STMSIH001654"  
 2004..2659  
 /gene="ribE"  
 /note="synonym: STM1426"  
 2004..2009  
 /gene="ribE"  
 /note="putative RBS for ribE; RegulonDB: STMSIH001655"  
 2018..2659  
 /gene="ribE"  
 /EC number="2.5.1.9"  
 /note="similar to E. coli riboflavin synthase, alpha chain (AAC74734.1); Blastp hit to AAC74734.1 (213 aa), 90% identity in aa 1 - 208"  
 /codon\_start=1  
 /transl\_table=11  
 /product="riboflavin synthase, alpha chain"  
 /protein\_id="AAL20348.1"  
 /db\_xref="GI:16419946"  
 /translation="MFTGIVQGTAKLVSIDKPNERTHVVTLPDYMLGLETGASVAH NGCCLTIVTEINGNISFDLMKLTITNLGALRVGDEVNVERAAKFSDEIGHLMSGH

gene  
 complement(2701..3860)  
 /gene="cfa"  
 /note="synonym: STM1427"  
 complement(2701..3849)  
 /gene="cfa"  
 /EC number="2.1.1.79"  
 /note="similar to E. coli cyclopropane fatty acyl phospholipid synthase (AAC74733.1); Blastp hit to AAC74733.1 (382 aa), 90% identity in aa 1 - 382"  
 /codon\_start=1  
 /transl\_table=11  
 /product="cyclopropane fatty acyl phospholipid synthase"  
 /protein\_id="AAL20349.1"  
 /db\_xref="GI:16419947"  
 /translation="MSSSCIEEVSVPDDNMYRIANELLSRADITINGSAPSDIRVKNP DFKRVLQEGSLGESYMDGNWECERLDIFFSKVLRAGLENPQHPHVKDTLRLGAR LNIQSKRAWIVGKEHYDLGNDLFSMLDPYMQYSICAYWKDADTLEAAQAKLKLIC EKLQOPGMRVLDTGCGWGLSQYMATHYGVSVGVTSIAEQKQAOCTRCGLDVSIL LEDYRLNDQDFRIVSVGMFHFVGPKNVTFVVDRLNLPKDGFLFLHTIGSKKTDHN VDPWKYIPFGNCLPSVROIABASESHFVMDWHNFADYDTTLMAWHERFINAWPE IAGNYNERFKRMFSYILNACAGAFRARDIQLWQVVFTEGVENGLRVRP"  
 complement(3855..3860)  
 /gene="cfa"  
 /note="putative RBS for cfa; RegulonDB: STMSIH001656"  
 complement(4139..5355)  
 /gene="ydhC"  
 /note="synonym: STM1428"  
 complement(4139..5344)  
 /gene="ydhC"  
 /note="similar to E. coli putative transport protein (AAC74732.1); Blastp hit to AAC74732.1 (403 aa), 85% identity in aa 1 - 403"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative MPS family transport protein"  
 /protein\_id="AAL20350.1"  
 /db\_xref="GI:16419948"  
 /translation="MQCKGFLVWLAGLSVLGFLATDMYLPAPAAIQADLOTAPAAVS ASLSFLAGFAVAQLMGSLDRYGRKPILLGLSIFALGSLGMLWBSAAALLTLRF VQAVGVCNATVWQALVTDYPSQKINRIFATIMPLVGLSPALAPILGSLWITLHFSWQ AIFALFVITULLMPLARLPKSPKARTEGQDKLTFAILLRSKYRGNVLIYACASAS FFALWTSPPFLTSAMGSPAVIGSLVVPQTAFILGIGYGCRAALQKWGQYLLPWLIG LYALSATWCGAGLLNNAISVEILIPFCVMAIANGAIYPIVVAQALRFPFQATGAA LQNTLQLCLGFLASLVSWLSITELLATTSVWLSITVLAALGYKMQSHADCAETGFPH ANVADKSH"  
 complement(5350..5355)  
 /gene="ydhC"  
 /note="putative RBS for ydhC; RegulonDB: STMSIH001657"  
 5447..6389  
 /gene="ydhB"  
 /note="synonym: STM1429"  
 5447..5452  
 /gene="ydhB"  
 /note="putative RBS for ydhB; RegulonDB: STMSIH001658"  
 5457..6389  
 /gene="ydhB"  
 /note="similar to E. coli putative transcriptional regulator LYSR-type (AAC74731.1); Blastp hit to AAC74731.1 (310 aa), 90% identity in aa 1 - 310"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative LysR family transcriptional regulator"  
 /protein\_id="AAL20351.1"  
 /db\_xref="GI:16419949"  
 /translation="MWSYSELVVDVAVNARNSFSAAOELHRVPSAVSVTVRQLEWL AVPLFVRRHRDVELTPAGVWFELKGRSVIKMQQTTRCQCIANGWRQALAVDNIV RPERTQMIVDFYRHFDELLVQFVFNQVWDALSDGRVELAIGAIQAIPIVGGYAF RDMGTLSWSCVVASDHPPLASMPGLSDTLNRNWPFLSREDVSRITPKRITWLLDNQKR VVVPDWESSATCLAGLCVGMVPTHFARQWIDSGKVALTLENPFDPFAACCVTWOONE ASPALWELDLVLDGSETLNRWLPREPAPDSGD"  
 complement(6386..7507)  
 gene

/genes="purR"  
/note="synonym: STM1430"  
complement (6386. .7411)  
/genes="purR"  
/note="purine nucleotide synthesis repressor."  
(SW:PURR\_SALT)  
/codon\_start=1  
/transl\_table=11  
/product="transcriptional repressor for pur regulon, glyA,  
glnB, prsA, speA (Galar/Laci family)"  
/protein\_id="AAL20352.1"  
/db\_xref="GI:16419950"

## CDS

## Alignment Scores:

Pred. No.: 6.82e-24 Length: 21913  
Score: 451.00 Matches: 91  
Percent Similarity: 64.94% Conservative: 22  
Best Local Similarity: 52.30% Mismatches: 49  
Query Match: 47.47% Indels: 12  
DB: 1 Gaps: 3

US-10-009-916A-1 (1-180) x AE008762 (1-21913)

QY 12 ValThrIleSerLeuThrSerIleThrSerValValLeuAlaCys----- 27  
Db 14525 TTAAGTTAGCGATGTGAGC-----CTGTGGCTGTGCGGTGGCGAG 14569  
QY 28 -----SerValThrSerGluValHisMetIleAspAspAsnGlyIleLysGlnSerIle 45  
Db 14570 GCGCCAGCGAGAAAGTAGAGTAATCTGTGACGGCGAAGCGTAGGACAGTCTATC 14629  
QY 46 GlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGly 65  
Db 14630 GGCACCGTCTCATCGATCAACCGAAGCGCGCTTAAATTTTACCCACACCTTAAAGCG 14689  
QY 66 LeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGlu 85  
Db 14690 TTGCCCGCGCGAGCATGCTTTTCATTCATTCATGCAACCGTAGCTGCCAGCGCGATT 14749  
QY 86 HisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisIleThrAspProAspLysThr 105  
Db 14750 AAAGACGGCAACCGGTTGCCGAGAGCGCTGGTGGTCACTGGACCCCAAAATACC 14809  
QY 106 GlyLysHisGluGlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValVal 125  
Db 14810 CGCAAGCATGAAGACCGGAGCGGCGCATCTCGCGACCTCCCGGTGTAGTCGT 14869  
QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLysGlu 144  
Db 14870 AATAATGATGGTATGCCGCGAACCCTGCTGCGCGCGCTCTGAAGTCTCTTGATGAA 14929  
QY 145 IleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLysProLeu 164  
Db 14930 GTGAAGATAAAGCGCTCATGATCCATGTGGCGGCGGATGAACATGTCGATCAGCCGAAA 14989  
QY 165 ProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
Db 14990 CCGCTCGCGCGCGGCGAATGCGTTACGCTGCGCGGTCAATT 15031

## RESULT 12

AL627271/c

## LOCUS

AL627271 Salmonella enterica serovar Typhi DNA linear BCT 04-JUL-2003

## DEFINITION

complete chromosome, segment 7/20.

## ACCESSION

AL627271 AL513382

## VERSION

AL627271.1 GI:16502733

## KEYWORDS

## SOURCE

## ORGANISM

Salmonella enterica subsp. enterica serovar Typhi

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Salmonella.

## REFERENCE

1 (bases 1 to 233050)

Authors Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D.,

Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T.G.,  
Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T.,  
Conerton, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N.,  
Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S.,  
Jagels, K., Krogh, A., Larsen, I.S., Leather, S., Moule, S., O'Goara, P.,  
Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J.,  
Stevens, K., Whitehead, S. and Barrett, B.G.  
Complete genome sequence of a multiple drug resistant Salmonella  
enterica serovar Typhi CT18  
Nature 413 (6858), 848-852 (2001)  
21534947  
11677608  
2 (bases 1 to 233050)  
Direct Submission  
Parkhill, J.  
Submitted (25-OCT-2001) Submitted on behalf of the Salmonella  
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,  
Hinxton, Cambridge CB10 1SA, UK  
E-mail: parkhill@sanger.ac.uk

## Notes:

Details of S. typhi sequencing at the Sanger Centre are available  
on the World Wide Web.  
(URL, [http://www.sanger.ac.uk/Projects/S\\_typhi/](http://www.sanger.ac.uk/Projects/S_typhi/)).

FEATURES  
Source

1. .233050  
Location/Qualifiers  
/organism="Salmonella enterica subsp. enterica serovar  
Typhi"

/mol\_type="genomic DNA"  
/strain="CT18"

/db\_xref="taxon:90370"  
139. .1581

/gene="STY1645"  
139. .1581

/gene="STY1645"  
139. .1581

/note="Similar to Pseudomonas aeruginosa  
arginine/ornithine antiporter arcD SW:ARCD\_PSPA (P18275)  
(482 aa) fasta scores: E(): 0, 45.5% id in 479 aa  
Fasta hit to YJDE\_ECOLI (445 aa), 31% identity in 450 aa  
overlap

Orthologue of E. coli ARCD\_ECOLI; Fasta hit to ARCD\_ECOLI  
(460 aa), 93% identity in 460 aa overlap

/codon\_start=1

/transl\_table=11

/product="putative amino acid permease"

/protein\_id="CAD01890.1"

/db\_xref="GI:16502734"

/db\_xref="GOR:98XGG1"

/db\_xref="TREMBL:Q8XGG1"

/translation="MEKLGLSALTALVLSMLGAGVFSLPQNMVAVASPAALLIGVA  
ITGVGILLIAFAMLITIRSELDGGIFTYAREGELTIGCSAWGYMLCAVIANVAT  
LVIVFSALSFDTTALRLFGDNTQWAIAGSVLLWIVHFLILRGVQTAANILVAT  
LAKLLPLGAFIVLAIMFKLDTLDTGVEIGIPVMBQVKNLTILTWVFGVAV  
VVSARKNRDVRATLAVLAALGILYLLVTLISGLVLAELAKRNPMSAGLWVVM  
MGPWEIITAGLIVSVCGAYLSWTIMAAEVFLAAAYSPFGEIARQNAQAPSASL  
WNTNCVQICLVLIWLTGSDYNTLTIASSEMILYFVLVGAFLKIATPLRAVGIG  
ACTYGLWLIYASGPMHLLSVVLYAPGLVFLYARRTHKHDNVLNRQEVVLTLGLIVA  
AVPATWMLVG"  
199. .1563

## misc\_feature

/gene="STY1645"

/note="Pfam match to entry PF00324 aa\_permeases, Amino  
acid permease, score -290.50, E-value 0.062"

complement (1643. .1978)

/gene="STY1646"

complement (1643. .1978)

/gene="STY1646"

/note="Orthologue of E. coli YDGC\_ECOLI; Fasta hit to  
YDGC\_ECOLI (111 aa), 81% identity in 111 aa overlap"

/codon\_start=1

/transl\_table=11

/product="putative membrane protein"

/protein\_id="CAD01891.1"

/db\_xref="GI:16502735"

/db\_xref="TREMBL:Q8Z6S2"

## gene

## CDS



```

/translation="MGLVTKAALGALVVVLIGLLSKTKNYYIAGLIPLFPTFALIAHY
IVASRGIDAMRTTIVSMNSLIPIFYIATLWYFSGVMRLPVALGAVVCGLSAWL
LIFCWIKWH"
2105..2836
/gene="STY1647"
/note="synonym: rsta"
2105..2836
/gene="STY1647"
/note="Similar to Staphylococcus aureus response regulator
YycF yycF TR:09XCM7 [EMBL:AF136709] (233 aa) fasta scores:
E(): 8.8e-28, 37.6% id in 234 aa
Fasta hit to TORR_ECOLI (230 aa), 31% identity in 213 aa
overlap
Fasta hit to BAER_ECOLI (240 aa), 34% identity in 229 aa
overlap
Fasta hit to YLCA_ECOLI (227 aa), 32% identity in 232 aa
overlap
Fasta hit to YEDW_ECOLI (239 aa), 34% identity in 233 aa
overlap
Fasta hit to OMPR_ECOLI (239 aa), 33% identity in 235 aa
overlap
Fasta hit to CPXR_ECOLI (232 aa), 37% identity in 233 aa
overlap
Orthologue of E. coli rsta (RSTA_ECOLI); Fasta hit to
RSTA_ECOLI (242 aa), 85% identity in 239 aa overlap"
/codon_start=1
/translation="putative two-component response regulator"
/product="putative two-component response regulator"
/protein_id="CAD01892.1"
/db_xref="GI:16502736"
/db_xref="GOA:Q8Z6S1"
/db_xref="TrEMBL:Q8Z6S1"
/translation="MNRVVFVEDDAEVGSLIAVLAKHDIIVIVEPRGDRADLILIT
OPDVLVDLMLPGKDGWTICRDLRHMGPVILTSLDSNMNHILALEMGACDYLKLT
TPPAVLARLRLHRSQEQTOQSLQESALTPKALRFGALITDPLNRAVQLNGDFI
SLSTADELLMELATHAGQIMDRDALLKRLGVNYDGLRSVDVAISRLRKLUDSAA
EPYRIKIRNKYGLIFAPHAWDETTC"
2108..2440
/gene="STY1647"
/note="pfam match to entry PF00072 response reg, Response
regulator receiver domain, score 62.70, E-value 7.9e-15"
2582..2794
/gene="STY1647"
/note="pfam match to entry PF00486 trans reg C,
Transcriptional regulatory protein, C terminal, score
62.90, E-value 1.8e-17"
complement(2901..3260)
/gene="STY1648"
complement(2901..3260)
/gene="STY1648"
/note="No significant database matches"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAD01893.1"
/db_xref="GI:16502737"
/db_xref="TrEMBL:Q8Z6S0"
/translation="MNORHPNLNIIISANANFSVMLLNTKVTIGNAVIYGEVVCYII
FMVVICISQFAKDGNTIAGRHVSINFWMLRVFFNLIFLSEKDVIMNRNAETVIK
GLFLILKIALMYTVSVF"
3317..4468
/gene="STY1649"
3317..4468
/gene="STY1649"
/note="Fasta hit to OMPC_ECOLI (367 aa), 66% identity in
383 aa overlap
Fasta hit to PHOE_ECOLI (351 aa), 61% identity in 386 aa
overlap
Fasta hit to YEDS_ECOLI (397 aa), 67% identity in 401 aa
overlap
Fasta hit to OMFP_ECOLI (362 aa), 58% identity in 392 aa
overlap
Fasta hit to NMPC_ECOLI (365 aa), 63% identity in 390 aa

```

```

overlap
Orthologue of E. coli OMPN_ECOLI; Fasta hit to OMPN_ECOLI
(377 aa), 80% identity in 388 aa overlap"
/codon_start=1
/transl_table=1
/product="outer membrane protein"
/protein_id="CAD01894.1"
/db_xref="GI:16502738"
/db_xref="GOA:Q56111"
/db_xref="Swiss-Prot:Q56111"
/translation="MKRKVLALVIPALLAAGAAHAELYNKGNKLDLYGKVDGLHYF
SDSSKDDQDTIMRVGFKGTQINDQITGQWYNYQANTTEGANSWTRALFAGL
KFGDYSFDYGRNRYVLYDVEGVTMLPEFGDSYTYADNTYMTGRANGVATYRTDFF
GLVDGLNALQYQKNSQSADDVNICTNNRNNGDDIRYDNGDGFSTYTDIGMGFS
AGAAATTSDRTNEQVAGGTAGDGADAWTAGLYDANNIYLATMYSETRMTPYGK
TDKGYDGVANKTQNFVTAQYQDFGLRPAVSFLMSKGLDLYNNVNGDDKDLVKYA
DVGATYFVNKNFSTYVDYKINLLDDDDPFYKDAGISDIDVALGNVYQF"
3395..4465
/gene="STY1649"
/note="pfam match to entry PF00267 Gram-ve porins, General
diffusion Gram-negative porins, score 619.80, E-value
6.8e-189"
4621..6327
/gene="STY1650"
4621..6327
/gene="STY1650"
/note="Similar to Mycobacterium tuberculosis hypothetical
57.3 kDa protein rv0552 or mtcy25d10.31 TR:006418
(EMBL:Z95558) (534 aa) fasta scores: E(): 1.8e-20, 26.4%
id in 561 aa"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="CAD01895.1"
/db_xref="GI:16502739"
/db_xref="TrEMBL:Q8Z6R9"
/translation="MLMISLVPLLSRTALLFLTATGAATAAPADIIHNGNIIT
INDAOPQASALAISSGRIVAIGDDTATNWRGDHRTIDLGKTVIPGLTTHHAIR
GGQWTFTETWYDPSLKDALDKLRADANRRPHDQWVAVGSIYPAQFAENRATVAE
LSHALPHPAVIOYLYVALVNGRIDVLGLNDTPPDLACIRYVERAKGATGKLFQ
DIAFNOLFASISSNADREGGLROFFADMNARGVTGIIDPSAGPAAAYEPYFAMNQG
DLPLRVGYRIPVQPEAKGHEQWFSNLMFAPARADDGQLAFLGLGESLVAGMDGVR

```

## Alignment Scores:

| Pred. No.:             | 9.38e-23 | Length:       | 233050 |
|------------------------|----------|---------------|--------|
| Score:                 | 449.00   | Matches:      | 91     |
| Percent Similarity:    | 64.94%   | Conservative: | 22     |
| Best Local Similarity: | 52.30%   | Mismatches:   | 49     |
| Query Match:           | 47.26%   | Indels:       | 12     |
| DB:                    | 1        | Gaps:         | 3      |

US-10-009-916A-1 (1-180) x AL627271 (1-233050)

| Qy | 12    | ValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCys                 | -----   | 27                            |
|----|-------|--|---|-------------------------------|
| Db | 37672 | TTAAGTTTAGCGTGGTGACG   | -----   | CTGCTGGCTGTGCGGTGCGGACG 37628 |
| Qy | 28    | -----  | SerValThrSerGluValHisMetIleAspAspAsnGlyIleIleYsglnSerIle 45 |                               |
| Db | 37627 | GCTGCCAGCGAGAAAGTAGATGAATCTGCTGACGCGCAAGCGGTAGGACAGTCTATC 37568  |   |                               |
| Qy | 46    | GlyThrValThrPheThrAspThrAspIleGlnIleLeuThrAspLeuIleYsgly 65      |   |                               |
| Db | 37567 | GSCACCGTCGTCATCGATGAACCGAGCGCGCTTAAATTTTACCCACACCTTAAACGC 37508  |   |                               |
| Qy | 66    | LeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGlu 85  |   |                               |
| Db | 37507 | TTCGCCCGCGGACGATGGTITTCACATTTCATGCCAAGTAGCTGCCAGCCCGCGATT 37448  |   |                               |
| Qy | 86    | HisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisIleThrAspProAspIle 105 |   |                               |
| Db | 37447 | AAAGACGCAAAAGCGGTTGCCGACAGCGCGTGTGTCTCTGATCTGACCAAAATACC 37388   |   |                               |





```
/locus tag="t1131"
complement(3058..3588)
/locus tag="t1131"
/note="Corresponds to STY1868 from Accession AL513382:
Salmonella typhi CT18"
/codon start=1
/transl table=11
/product="putative cytochrome"
/protein id="AAO68792.1"
/db_xref="GI:29137229"
/translation="MAHFSRLQITLHWLTLTLTGIAVAAIELRWAPKGSVYLFMKD
MYDMGLVWALMELRLYLKHYPDPVITPPPHQHVAAKLMLHIALYLTFLALPLIG
VAMASGCKWSFFGFTVPVFLTPDSTLKSIDIKRIHEMLANIGYFLIAMHAAALFHH
YIQKDTFSRMLPGKS"
complement(3737..4051)
/locus tag="t1132"
/codon start=1
/transl table=11
/product="putative lipoprotein"
/protein id="AAO68793.1"
/db_xref="GI:29137230"
/translation="MTNKKHIFSIIFIGSLITGCATGSPPTGIGLYTDVKGPITATSL
PATKTGACATVLTGIVNTGDISAKKAGDISLVSSVDYETTGSPFFYKGTCTVVR
GQ"
4755..6389
/locus tag="t1133"
/locus tag="t1133"
/note="Corresponds to STY1865 from Accession AL513382:
Salmonella typhi CT18"
/codon start=1
/transl table=11
/product="putative substrate-binding transport protein"
/protein id="AAO68794.1"
/db_xref="GI:29137231"
/translation="WLRKQKORIIMDKLMKGLALVTCALTIVFTSSLFAASDTA
DGRTLKLAIGPEPTGFPDMLGWSHGSYLLHAPLLKQNDMSWGLDLTEKVDITSPDG
KIWTLKPLGKFSGLPTAEDVVFTYNTKAAKSGSKDKIDMGNFSHARALDRIRMTLL
SHQSTFNVYLVGLSGLVIFASRYDEKTFAREPIGAPRYLRFVQFQQLIVRANPWYAG
KPNQDNRLVFLVLDNDAVAAARSGQLGLVRIAPSMVTPQDNLKLVWRDSVENRGI
VFPMPYACKDANGPVGNVDVAVIRAINYAINRKOLAEQVMEGHAI PAYSAVQGI
LEWQNPVSFVSDGDIKARATILEAGWKINSAGVREKAGKEARLTLVASGDSRRLD
AEVNRAMLQRGIVVLSQSGSWETVERHMHANPTLFGWGLDPMELFHHYSGKAGVE
YNYGYINPAVEAHLKAIDAPWQKAIPFQVQVWDKQAGVQGDAAWALLNIG
HTYLANPCIDLGKSPETHGWSVLNLDLDDWTCTCR"
6386..7363
/locus tag="t1135"
/locus tag="t1135"
/note="Corresponds to STY1864 from Accession AL513382:
Salmonella typhi CT18"
/codon start=1
/transl table=11
/product="putative inner membrane transport protein"
/protein id="AAO68795.1"
/db_xref="GI:29137232"
/translation="MMKLFFSHFHLRIILLVVAAGTFILLSPSPVPIRATVIGNDLL
HYPPQYARIARWGLDQPLWRFHGFHFWFRLQDGYSMLFNMPVASVIRERFATSF
ALLAGATLISGLVGLFAGFLRFPWDMICRISYLLSLPTFTWIMLLALFVAV
RNPVLPCVWCPDGNAGTALLSERLHLVPCVCLSLGMSQIALHTRKTIASVMKS
EPIRARQAGDKMSILSHOVLRAHTPALCLQFASLGLMGALLAEKVPAYPGLQG
ATIDAGLRDGVFLMGLVILFCTLLVPAGNTISAMLVVNLRSRLRPDAL"
7353..8165
/locus tag="t1134"
/locus tag="t1134"
/note="Corresponds to STY1863 from Accession AL513382:
Salmonella typhi CT18"
/locus tag="t1134"
/note="Corresponds to STY1863 from Accession AL513382:
Salmonella typhi CT18"

Alignment Scores:
Pred. No.: 1.2e-22 Length: 300523
Score: 449.00 Matches: 91
Percent Similarity: 64.94% Conservative: 22
Best Local Similarity: 52.30% Mismatches: 49
Query Match: 47.26% Indels: 12
DB: 1 Gaps: 3

US-10-009-916A-1 (1-180) x AE016838 (1-300523)
QY 12 ValThrLeuSerLeuLeuThrSerLeuValValLeuAlaCys----- 27
Db 169329 TTAAGTTTAGCATGAGTGGTGACG-----CTGCTGGCTGTGCGGGTGGCGAG 169373
QY 28 -----SerValThrSerGluValHisMetIleAspAspAspGlyIleLeuGlnSerIle 45
Db 169374 GCTGCCAGCGAGAAAGTAGATGAATCTGTCAGCGCGGCGGAGGACAGTCTATC 169333
QY 46 GlyThrValThrPheThrAspThrAspGlyLeuGlnIleLeuThrAspLeuGly 65
Db 169034 GGCACCGCTGCTCATCGATGAACCGAGCGGCTTAAATTTACCCACACCTTTAAAGCG 169093
QY 56 LeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGlu 85
Db 169094 TTGCCCGCGGCGAGCATGGTTTTCACATTCATGCCAAAGTAGCTGCCACCGCGGATT 169153
QY 86 HisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisIleTyrAspProAspLysThr 105
Db 169154 AAGACGCGCAAGCGGTGCCGAGAACCGCTGCTGTCATCTGGACCCACAAATACC 169213
QY 106 GlyLysHisGluGlyProLeuGlyAsnGlyHisIleGlyAspLeuProArgLeuValVal 125
Db 169214 GGTAAAGCATGAAGACCGGAGGTCAGGGCATCTGGCGGACCTCCCGGTGTAGTCGTT 169273
QY 126 LysAlaAspGlyIleAlaLysGlyThrLeuLeuAlaProArgLeu---ThrValLysGlu 144
Db 169274 AATATGATGATATCGCCACCGGTTACTGGCGCGCGCTGGAAGTCTCTGCATGAA 169333
QY 145 IleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLysProLeu 164
Db 169334 GTGAAGATAAAGCGCTCATCATCATGTGGCGGCGGATACATGTCGATCAGCGGAAA 169393
QY 165 ProLeuGlyGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db 169394 CCGTCCGCGCGCGGCGGAAACGCTTACGCTCGCGCGCTCAT 169435

RESULT 14
AE016860
LOCUS
DEFINITION
Pseudomonas syringae pv. tomato str. DC3000 section 5 of 21 of the
complete genome.
ACCESSION
AE016860 AE016853
VERSION
AE016860.1 GI:28851610
KEYWORDS
Pseudomonas syringae pv. tomato str. DC3000
SOURCE
Pseudomonas syringae pv. tomato str. DC3000
ORGANISM
Pseudomonas syringae pv. tomato str. DC3000
REFERENCE
1 (bases 1 to 311249)
AUTHORS
Buell C.R., Joardar V., Lindeberg M., Selenberg J., Paulsen I.T.,
Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
Nelson W.C., Davidsten T., Zafar N., Zhou L., Liu J., Yuan Q.,
Khouri H., Fedorova N., Tran B., Russell D., Berry K.,
Kutterback T., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
Deng W.L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
Telaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J.,
Tang X., Bender C.L., White O., Fraser C.M. and Collier A.
The complete genome sequence of the Arabidopsis and tomato pathogen
Pseudomonas syringae pv. tomato DC3000
Proc. Natl. Acad. Sci. U.S.A. 100 (18), 10181-10186 (2003)
12928499
PUBMED
REFERENCE
2 (bases 1 to 311249)
```

## AUTHORS

Buell, R., Joardar, V., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Utterback, T., Van Aken, S., Feldblyum, T., Gwinn, M., Dodson, R., DeBoy, R., Durkin, A., Kolonay, J., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M., Haft, D., Selengut, J., Nelson, W., Davidsen, T., White, O., Fraser, C. and Collier, A.

## TITLE

Submitted (03-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

## JOURNAL

Location/Qualifiers

gene

1. .311249

## FEATURES

source

/organism="Pseudomonas syringae pv. tomato str. DC3000"

/mol\_type="genomic DNA"

/strain="DC3000"

/db\_xref="taxon:223283"

/note="pathovar: tomato"

162. .818

/locus\_tag="PSPT01158"

162. .818

/locus\_tag="PSPT01158"

/note="identified by match to PFAM protein family HMM PF00440"

/codon\_start=1

/transl\_table=11

/product="transcriptional regulator, Tetr family"

/protein\_id="AA054687.1"

/db\_xref="GI:28851611"

/translation="MVKTTREPSATSLKRRMRMEKRTLLDAALEIFSRVGVHSS

LQVSLADVSTNLVYFSSKDLVNLVRLVLSPLVHFTADKEPQVQVIGYI

KAKLEMSRDPHAEISRLFCNEVMQGLQGLQHPDRVQAKVAVIQHWISQIAP

INPHLIFTLMWATQHYADTQVEAVTGKILDDPVPFBEVLASRSMVLDGILPRTA

"

1052. .2143

/locus\_tag="PSPT01159"

1052. .2143

/locus\_tag="PSPT01159"

/note="similar to GP:13937517; identified by sequence

similarity; putative"

/codon\_start=1

/transl\_table=11

/product="bmp family protein"

/protein\_id="AA054688.1"

/db\_xref="GI:28851612"

/translation="MYKKLLPGTLCAAVGAAGFAQAAGRLTKGPAKIMVYISPRN

DGWTQAFDEARLEKLDLTKTFVESVPEAAATPVVDRLIARGANVIVGTAFY

STFLALAKKYPDVAFPLNGSTTNAANLESFYGRVTSOYVLCMAAGAASTKGLGFV

AAAPGQVNTTNAVELGAROTNBKATVIVTGAWDPVKERAATMALIDNGDVVG

QHVDSPTQIVAQERGIHGTGHRDLSEFAPKATVCSVVWDRFLGPKLKIAGNW

TPAANGALLSMQGGDDITLTGDTPIISAENRKKIEDERAELLAGKKIVYSGFLADRDG

KERVAAGQQLSDPDLWMKDMFVEGVKTOQ"

2140. .3651

/locus\_tag="PSPT01160"

2140. .3651

/locus\_tag="PSPT01160"

/note="similar to GP:15023582; identified by sequence

similarity; putative"

/codon\_start=1

/transl\_table=11

/product="ABC transporter, ATP-binding protein"

/protein\_id="AA054689.1"

/db\_xref="GI:28851613"

/translation="MSQVNAQLQIGISFSGDKFALIDADFTACMGVEHALGENGAG

KSLNMIAGLYAPETGLMDNDVQLQPKDASRYIGMVHQFPLKVPPTVAQNI

LLALPEQCEGYRKRRLMLERDVRKAELGFAIDPSRVVDSLSVAEQRVIELKVL

LAGARILLIDEPVAVLTDQEAERLLITVQAFARQAGNAVILVTHKADVRYADRTVM

RGRTIOTLDPQSVSVQVLTUVGESVPAAHFPAVPGVEVRLQVLDLRSVGSALNG

VNMTLRAGIYGIAGVGGQAEALNALMGLPQATEGTHLAPFGDLRNASAEQRQL

RTIAPADRYGAALSGLSVAENFGINTHSYGYSFWRLGYKRLQDQARAVESFDV

QVRSIDQKAAALLSGNAQKLVIAREFSRDPQLVLVHSPSRGLDVRAVQADHRLRAA

ROAGAAVLLISDDLDELVALADRIQVWNGGRIVAEFDHPADRAQKAVWVSHD"

3644. .3838

/locus\_tag="PSPT01161"

3644. .3838

/locus\_tag="PSPT01161"

/note="identified by Glimmer2; putative"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein"

/protein\_id="AA054690.1"

/db\_xref="GI:28851614"

/translation="MTELAQPLLPITARQPLLRYYTLELRQQTGWPLQALIIALA

LVVGLAICTAIIILAGVPA"

3854. .4764

/locus\_tag="PSPT01162"

3854. .4764

/note="this region contains a gene with one or more

premature stops or frameshifts, and is not the result of a

sequencing artifact; ABC transporter, permease protein,

degenerate; similar to GP:15139998; identified by sequence

similarity; putative"

4764. .5690

/locus\_tag="PSPT01163"

4764. .5690

/note="this region contains an authentic point mutation,

causing a premature stop, and is not the result of a

sequencing artifact; ABC transporter, permease protein,

authentic point mutation"

5804. .6493

/locus\_tag="PSPT01164"

5804. .6493

/locus\_tag="PSPT01164"

/note="similar to GP:5123490; identified by sequence

similarity; putative"

/codon\_start=1

/transl\_table=11

/product="ompA family protein"

/protein\_id="AA054691.1"

/db\_xref="GI:28851615"

/translation="METSRLIIVATAVALLSGCATSNPDNQGQASGSGMSKTAKY

GGGLAGNAVAGAAIDHNRRKGALIGAVAGAASAGYGYADKQALRASMANVTGV

EVQRQDQIKLIMPENITPATDSAIASSFYSPLANNLSKQFNQNNIIEIIGYTST

GRQNMNLSQQRAQSVATYLTSGQVDSHLSVRGAGPDQPIASNADANGRAQNRRE

VNLKPIPGQQ"

6822. .7922

/locus\_tag="PSPT01165"

6822. .7922

/locus\_tag="PSPT01165"

/note="identified by Glimmer2; putative"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein"

/protein\_id="AA054692.1"

/db\_xref="GI:28851616"

/translation="MFFHMLPSWDSIRRLQDRDIAPYRGHFTPRNEPVPSPALIM

RRFNCVSAFDRAEWEACRILQRFVDLDSIIKDLIDVVTQMAMIVVGSFVGGAA

GAGAGVFGVGAVPGLMGALGAQASAFILGLEIAESVIDGPPRIETTYTRG

ITTAHQGPBEOGNPFGMDLLAQAQNSAQVEIAQGHVEVILLIGAMVQVITRGNAR

ALANEMRASAKERLQWMLHEDEIKRPPDQRPPEPRGALAEQETLSTNRKATADA

AKPDAMDQKTVSSKIDPKQLDLVTHGPNHLSLSVHGGKVSQDGLMLRAKSAIKPNDGV

GPSVPLSSAPHSDELLIFADQITRETLKKS"

7963. .8940

/locus\_tag="PSPT01166"

7963. .8940

/locus\_tag="PSPT01166"

/note="similar to GP:13094159, and GP:4514540; identified

by sequence similarity; putative"

/codon\_start=1

/transl\_table=11

/product="ISPSey, transposase"

/protein\_id="AA054693.1"

/db\_xref="GI:28851617"

/translation="MKQMTFADAAYAGKRKHTRRFRFLIEMDQVVPVWKGILALIBPHY

PKGEGEPAYPLMAMLRTHLMQNWFGYSDPAMEEALYETTLRQFAGSLDRADETT

INFERRLLEKELAGLILQVTLNGYLDGRLMROGVVDARTIHPSSTKWKGKEDP

EMHQTKGNQVFFGKKAHIGVDAESGLVSLVGTAAVADVTVQVOLLHGERTYVSGD

AGYTVGDKRAEHQDMQMIWSTAAPSRYYKHGKSLIARVYRKIEFTLKAQRAKVEHP

FRVYKQFGYTKVFRFRGLAKNTAQOATLFLALNSLMMVRKRLIANGEVRL"

10294. .10887

/locus\_tag="PSPT01167"



```

DEFINITION Sequence 2340 from patent US 6610836.
ACCESSION AR385611
VERSION AR385611.1 GI:40095345
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 435)
AUTHORS Breton,G.L. and Osborne,M.
TITLE Nucleic acid amino acid sequences relating to Klebsiella pneumoniae
for diagnostics and therapeutics
JOURNAL Patent: US 6610836-A 2340 26-AUG-2003;
FEATURES
    source
        Location/Qualifiers
            1..435
                /organism="unknown"
                /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 6,53e-24 Length: 435
Score: 429.50 Matches: 81
Percent Similarity: 70.92% Conservative: 19
Best Local Similarity: 57.45% Mismatches: 39
Query Match: 45.21% Indels: 2
DB: 6 Gaps: 1
US-10-009-916A-1 (1-180) x AR385611 (1-435)
QY 40 GYllelysglnserilegylthrvalthrphethrAspThr-AspLysGlyLeuGlnl 59
Db 9 GCGCTGGTCAGACCATCGCAGCGTCAAAATCACCAGAAACCGCGCGGACTCGAGTT 68
QY 59 elysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyG 79
Db 69 CGCCCCCATCTGCGGGCGCTACCGCGGTAGACAGCGGTTTCATATTCATGCCAAGG 128
QY 79 ySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHi 99
Db 129 CAGCTGCCAGCGCGCGATGAAGAAGTAAAGCGCGCGCGCGCGCGCGGCGGAC 188
QY 99 sTyrAspProAspLysThrGlylyshisGluGlyProLeuGlyAsnGlyHisLysGlyAs 119
Db 189 TTACGATCGCAGCATACCGGCAAAACAGCAAGGCGCGTTGGGGCGCGGCGATCTTGGCG 248
QY 119 pLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProAr 139
Db 249 CTTGCCCTCTGCTGGTGTCAACATCGCGCGGTAGCGGACCGCATTTATGTCICGCG 308
QY 139 gLeu---ThrValLysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAspAs 158
Db 309 CCTGAAACGCTGGCGAGGTGAAGAGAAAGCGCTGATGTCACGTAGCGGGGATAA 368
QY 158 nTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValll 178
Db 369 CATGCCGACAGCCCGACCGCTGGCGCGCGCGCGCAACCGTTTTCCTCGCGGGTGAT 428
QY 178 e 178
Db 429 T 429
RESULT 17
STCZSODC
LOCUS Salmonella typhimurium sodC gene.
DEFINITION
ACCESSION Y13121
VERSION Y13121.1 GI:2462698
KEYWORDS copper/zinc-superoxide dismutase; sodC gene.
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE 1
AUTHORS Farrant,J.L., Sansone,A., Canvin,J.R., Pallen,M.J., Langford,P.R.,

```

```

Wallis,T.S., Dougan,G. and Kroll,J.S.
Bacterial copper- and zinc-cofactored superoxide dismutase
contributes to the pathogenesis of systemic salmonellosis
Unpublished
2 (bases 1 to 696)
Kroll,J.S.
Direct Submission
Submitted (12-MAY-1997) J.S. Kroll, Imperial College School of
Medicine, Dept. of Paediatrics, St Mary's Hospital, Norfolk Place,
London, W2 1PG, UK
FEATURES
    source
        Location/Qualifiers
            1..696
                /organism="Salmonella typhimurium"
                /mol_type="genomic DNA"
                /strain="ST4/74"
                /db_xref="taxon:602"
            83..616
                /gene="sodC"
            83..616
                /gene="sodC"
            /EC number="1.15.1.1"
            /codon_start=1
            /transl_table=11
            /product="copper/zinc-superoxide dismutase"
            /protein_id="CAA73588.1"
            /db_xref="GI:2462699"
            /db_xref="GOA:P53636"
            /db_xref="Swiss-Prot:P53636"
            /translation="MKYITLLVAGALISCSMAENTLTVKNDALSSGTGNGIGBIT
VSETPYGLFLFTPLHGLTEGHHGHPNFCMPGMDKGEVPAALMAGHLDPEKTK
HLGPYNDKHLGLDPLVNVADGTATYPLLPRLKSLSELKGLHSLMHKGDYSDKP
APLGGGARFACGVIEK"
            659..683
                terminator
            /evidence=not_experimental
ORIGIN
Alignment Scores:
Pred. No.: 1.88e-23 Length: 696
Score: 426.00 Matches: 88
Percent Similarity: 63.64% Conservative: 24
Best Local Similarity: 50.00% Mismatches: 56
Query Match: 44.84% Indels: 8
DB: 1 Gaps: 4
US-10-009-916A-1 (1-180) x STCZSODC (1-696)
QY 11 IleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerValThr 30
Db 80 CTAATGAATAACACATATTCGCTGTGAGTGTGCTCATCATCATGTTGTTTCAGCAATG 139
QY 31 SerGlu-----ValHisMetIleAsp-----AspAsnGlyIleLysGlnSer 44
Db 140 GCAGAGATACCTGACTGTAAGATGAACGATGCCCTCTCCAGCGGAACAGAGAAAC 199
QY 45 IleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLys 64
Db 200 ATAGGTGAATCACAGTTTCAGAGACACCTTACGGTCTCTCTTTCACCTCCTCAATAAT 259
QY 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAla 84
Db 260 GGTCTTACGCCAGGAATTCACGGCTTCCATGTCTCCACACAAACCAAGTGTATGCCGGA 319
QY 85 GluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLys 104
Db 320 ATGAAGACGGTAAAGAGGTTCCGCGCTCATGGCGGAGGACATCTTACCCCGGAAA 379
QY 105 ThrGlyLysHisGluGlyProLeuGlyAsn---GlyHisLysGlyAspLeuProArgLeu 123
Db 380 ACCGGGAACATCTTGGCCCATATATGACAAAGGCGCATTTGGGGATCTGCTGGACTG 439
QY 124 ValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeu---ThrVal 142
Db 440 GTTGCAATGCAGATGGTACAGCCAGTATCGGTTACTGGCACCACCGCCTTAATCACTG 499

```

```

Qy 143 LysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLys 162
Db 500 TCAGACTGAAGGTCACCTCAATGATGATCATTAAGGCGGTGACATTAATCCGATANA 559

Qy 163 ProLeuProLeuGlyGlyGlyGlyValaArgIleAlaCysGlyValIle 178
Db 560 CCTGCTCCACTGGGTGGTGGCGGTGCACGTTTGGCTGTGGTGCATT 607

RESULT 18
AJ620903 840 bp DNA linear BCT 01-MAY-2004
LOCUS Salmonella enterica subsp. enterica serovar Choleraesuis sodC1 gene
DEFINITION for superoxide dismutase [Cu-Zn] precursor.
ACCESSION AJ620903.1 GI:46934750
VERSION sodC1 gene; superoxide dismutase [Cu-Zn] precursor.
KEYWORDS Salmonella enterica subsp. enterica serovar Choleraesuis
SOURCE Salmonella enterica subsp. enterica serovar Choleraesuis
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE 1 Amendola,S., Ajello,M., Pasquali,P., Kroll,J.S., Langford,P.R.,
AUTHORS Rotilio,G., Valenti,P. and Battistoni,A.
TITLE Differential contribution of sodC1 and sodC2 to intracellular
survival and pathogenicity of Salmonella enterica subsp. enterica
serovar Choleraesuis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 840)
AUTHORS Battistoni,A.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2004) Battistoni A., Biology, University of Rome
Tor Vergata, via della Ricerca scientifica, 00133, ITALY
FEATURES
source
1..840
/organism="Salmonella enterica subsp. enterica serovar
Choleraesuis"
/mol_type="genomic DNA"
/serovar="Choleraesuis"
/isolate="A50"
/sub_species="enterica"
/db_xref="taxon:119912"
/country="United Kingdom"
193..726
/gene="sodC1"
193..726
/gene="sodC1"
/EC_number="1.15.1.1"
/function="disproportionation of superoxide into hydrogen
peroxide and molecular oxygen"
/codon_start=1
/transl_table=11
/product="superoxide dismutase [Cu-Zn] precursor"
/protein_id="CAF06531.1"
/db_xref="GI:46934751"
/translacion="MKYITILSVAGALISCSAMAEINTLVKMDALSSGTGENIGEIT
VSETPYGLLPFLHNGLPFGTHGFVHTNPSCPMKDKGEVPALMAGCHLDPEKTGK
HLGPVNDKGHGLDPLGLVNVADGTATYLLAPRLKSLSELKSHLMHKGNDYSDKP
APLGGGARFACGVIEK"
193..726
/gene="sodC1"
253..723
/gene="sodC2"
/product="superoxide dismutase [Cu-Zn]"
253..723
/gene="sodC2"

sig_peptide
mat_peptide
gene
ORIGIN
Alignment Scores:
Pred. No.: 2,25e-23 Length: 840
Score: 426.00 Matches: 88
Percent Similarity: 53.64% Conservative: 24
Best Local Similarity: 50.00% Mismatches: 56

```

```

Query Match: 44.84% Indels: 8
DB: 1 Gaps: 4
US-10-009-916A-1 (1-180) x AJ620903 (1-840)

Qy 11 IleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerValThr 30
Db 190 CTAATGAATAACACAATATTTCGTCGTAGCTGGTGGCTCATCAGTTGTTCAATATG 249

Qy 31 SerGlu-----ValHisMetIleAsp-----AspAsnGlyIleIleGlnSer 44
Db 250 GCAGAGAATACCCCTGACTGTAAGATGAACATGCCCTGTCCAGCGGACACAGGAAAAAC 309

Qy 45 IleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLys 64
Db 310 ATAGGTGAATCACAAGTTTCAGACACACCTTACGGTCTGCTTTTCACCTCCCTCACTAAT 369

Qy 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAla 84
Db 370 GGTCTTACGCCAGGAATTCACGGCTTCATGCCACACAAACCCCAAGTTGTATGCCGGA 429

Qy 85 GluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLys 104
Db 430 ATGAAGAAGCGTAAGAGGTTCCGGCGCTCATGGCCGAGGACATCTTGACCCCGAAAAA 489

Qy 105 ThrGlyLysHisGluGlyProLeuGlyAsn---GlyHisLysGlyAspLeuProArgLeu 123
Db 490 ACCGGAAACATCTTGGCCATATAATGACAAAGGCATTTGGGGGATCTGCCTGACTG 549

Qy 124 ValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeu---ThrVal 142
Db 550 GTTGTCAATGAGATGGTACAGCCACGCTATCGCTTACTGGCACCACGCTTAAATCACTG 609

Qy 143 LysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLys 162
Db 610 TCAGAACTGAAGGTCACCTCAATGATGATCATTAAGGCGGTGACATTAATCTCGATAAA 669

Qy 163 ProLeuProLeuGlyGlyGlyGlyValaArgIleAlaCysGlyValIle 178
Db 670 CCTGCTCCACTGGGTGGTGGCGGTGCACGTTTGGCTGTGGTGCATT 717

RESULT 19
AJ608743/c 62027 bp DNA linear BCT 23-APR-2003
LOCUS Salmonella typhimurium LT2, section 49 of 220 of the complete
DEFINITION genome.
ACCESSION AE008743 AE006468 AE008744 AE008745
VERSION AE008743.2 GI:22024639
KEYWORDS Salmonella typhimurium LT2
SOURCE Salmonella typhimurium LT2
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE 1 (bases 1 to 62027)
AUTHORS McClelland,M., Sanderson,K.E., Spieth,J., Clifton,S.W.,
Latreille,P., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F.,
Hou,S., Layman,D., Leonard,S., Nguyen,C., Scott,K., Holmes,A.,
Grewal,N., Mulvaney,E., Ryan,B., Sun,H., Florea,L., Miller,W.,
Stoneking,T., Nhan,M., Waterston,R. and Wilson,R.K.
TITLE Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2
JOURNAL Nature 413 (6958), 852-856 (2001)
MEDLINE 21534948
PUBMED 11677609
REFERENCE 2 (bases 1 to 62027)
AUTHORS The Salmonella typhimurium Genome Sequencing Project
CONSRTM Direct Submission
TITLE Submitted (29-MAR-2001) Genome Sequencing Center, Department of
JOURNAL Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
COMMENT On or before Jul 31, 2002 this sequence version replaced
gi:16419526, gi:16419555, gi:16419507.

```

COMMENT Supported by NIH grant 5U 01 AI43283

Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs: GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and GeneMark; <http://opal.biology.gatech.edu/GeneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc; <http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and RegunonDB; [http://kinich.cifn.unam.mx:8850/db/regulondb\\_intro.frameset](http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset)

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

#### FEATURES

##### source

Location/Qualifiers  
 join(1..13417.58905..62027)  
 /organism="Salmonella typhimurium LT2"  
 /mol\_type="genomic DNA"  
 /strain="LT2; SGSC 1412; ATCC 700720"  
 /db\_xref="ATCC:700720"  
 /db\_xref="taxon:99287"  
 /focus

##### source

13418..58904  
 /organism="Phage Gifsy-2"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:129862"  
 137..2051

##### gene

/gene="ycbB"  
 /note="synonym: STM0995"

##### RBS

137..202

/gene="ycbB"

/note="putative RBS for ycbB; RegulonDB:STMSIH001279"

210..2051

/gene="ycbB"

/note="similar to E. coli putative amidase (AAC74011.1); Blastp hit to AAC74011.1 (615 aa), 84% identity in aa 1 - 615"

##### CDS

/codon\_start=1

/transl\_table=11

/product="putative periplasmic protein"

/protein\_id="AAL1929.1"

/db\_xref="GI:16419508"

/translation="MLNKKRCGRCLSAISLCIAVTFAPLFNAQAQDEPEMIPGDSAVAA  
 TDLAQKQSAATAMAGIQPLPEGSAEKVRADLQSLPGSYTPVYMSQLTLLVAAR  
 DKPMWNRDVAQQAQAEVAIAIQFQFTAWALLDPAVGMARDVLAQVSDAMQ  
 YLHFTANIIVKQRMLYSKNPALATPPVSVINQIOALEGOLPMFVASLAPQHPQY  
 APMDALLKVADSPWCLTNTATLRPGQMSNDVPALEIILQRTGMLDGGPKIALPG  
 DNTADSAVVSPSAVDETSVAHDEPTARSKAPARAYDRELVEAVKSPQAWQGLGA  
 DGVIGPATKLNMTPAQAGVIALNIQRLRLPDELSTGIMWNPAYSLVYIQNGNQ  
 VLASRVIGRPDRKTPMMSALNNVNPVNPPTLARKDILPKVMNDPGLYERHGY  
 TPNHTLIFQDARALSGCGVRNKAISELANMLQDAGNDARISGALKQGDTRVNVIRQ  
 NIPVNLVYLTAFIGADGRQYRTDILVNYDLTARSSAQIVPKVEQLIR"  
 2311..2859

##### gene

/gene="ycbB"

/note="synonym: STM0996"

2311..2859

/gene="ycbB"

/note="similar to E. coli orf, hypothetical protein (AAC74012.1); Blastp hit to AAC74012.1 (182 aa), 95% identity in aa 1 - 182"  
 /codon\_start=1  
 /transl\_table=11

/product="putative outer membrane protein"  
 /protein\_id="AAL19930.1"  
 /db\_xref="GI:16419509"  
 /translation="MDKFDANRRKLLALGGVALGAAILPAPAPAFATLSTPRPRILLTNN  
 LHTGSIKAEFDGGRAYIQDELAKUNHFRDYRANKVRSIDPRLPOLYRLQGLGTR  
 KVFQISIGYRSLDTNNELFARSQVAKKSYHTKQAMDHFIEGVALSNIRKKAALSRA  
 CGVGYPKRSNPFVHIDTGPAPHW"  
 2877..3534  
 /gene="ycbL"  
 /note="synonym: STM0997"  
 2877..2882  
 /gene="ycbL"  
 /note="putative RBS for ycbL; RegulonDB:STMSIH001280"  
 2887..3534  
 /gene="ycbL"  
 /note="similar to E. coli orf, hypothetical protein (AAC74013.1); Blastp hit to AAC74013.1 (215 aa), 91% identity in aa 1 - 215"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative metallo-beta-lactamase"  
 /protein\_id="AAL19931.1"  
 /db\_xref="GI:16419510"  
 /translation="MNYRIIPVTFNSQNSLIWCQTRLAALVDPGDAEKIKQEVDA  
 SGVILMQILLTHGLDHYGAASELAQHVGPVIGPEKEDEFWLQGLPADSRMFGLEDC  
 QPLTPDRMLNDGRYSVGNVTLOVLHCPGHTGHVVFDEQSLISGVIRPKGVGR  
 SDFPRGDHTQLJDAIKRKLPLGDDVTPIFGHGPLSTIGYERLHNPFLQDEMPVW"  
 complement(3596..4799)  
 /gene="aspC"  
 /note="synonym: STM0998"  
 complement(3596..4786)  
 /gene="aspC"  
 /EC\_number="2.6.1.1"  
 /note="similar to E. coli aspartate aminotransferase (AAC74014.1); Blastp hit to AAC74014.1 (396 aa), 95% identity in aa 1 - 396"  
 /codon\_start=1  
 /transl\_table=11  
 /product="aspartate aminotransferase"  
 /protein\_id="AAL19932.1"  
 /db\_xref="GI:16419511"  
 /translation="MFENITAAPADPILGLADLFRADRRPGKINLIGVYKDETGTPT  
 VLTSVKABOYLLENNTKYLIGIDIPBFARCTOELLFGKGSALINDYKARTAPQPG  
 GTGARLAAADFLKNTPVKRVWSPNSPNHKSVENAGLEVYVAYDAENHTLDPE  
 ALQASLSEAQAGDVILFHGCCNPTGIDPTLQMQVLAELSVEKWLPLDFEAYQGF  
 RGEEDAEGRAPAAHLKELIVASSYKNGFYNERVAGCTLVADAETVDFRAFSOMK  
 SAIRAYSNPPAHGASIVATILSNDALRAIWEQELTDMQRIOIRMQLFVNTLQEKGA  
 NRDFSIIKQNGMFSFGLTKDQVLRLEEFVGYAVASGRVNVAGTPTDNPAPLCEAI  
 VAVL"  
 complement(4794..4799)  
 /gene="aspC"  
 /note="putative RBS for aspC; RegulonDB:STMSIH001281"  
 complement(4971..6162)  
 /gene="ompF"  
 /note="synonym: STM0999"  
 complement(4971..6062)  
 /gene="ompF"  
 /note="outer membrane protein F precursor (PORIN OMPF)  
 (outer membrane protein 1A, 1A, or B). (SW:OMPF\_SALTY)"  
 /codon\_start=1  
 /transl\_table=11  
 /product="outer membrane protein 1a (1a;b,f), porin"  
 /protein\_id="AAL19933.1"  
 /db\_xref="GI:16419512"  
 /translation="MMKREKILAAVLPALLAAATANAAYNNKDGKLDLYGKAVGRHV  
 WTTTGDGSKNADQYAIQFGKGTQINTDLTGQWEYRTKADRAEGEQNSNLVRLAF  
 AGLKYAEVSDYDGRNIGYVDVDSYTPMAPYFTWGGAYTDNMTSRAAGLLTYR  
 NSDFELVDLSFGIQYQKQDNHSINSQNGDGVGYTMAVEFDGVTYAAVNSKRT  
 NDQDRDNGDRAESWAGAKIDANNVLAAYASTRNMSIVENTVITVENANKIQN  
 LEVVAQYDFDLRPAISYQSKGQLNGAGSADLAKYIQAGATYTFNNKNNVWDY  
 RFNLADYDSSSYVTGDOAAVGYTQF"  
 complement(6071..6076)  
 /gene="ompF"

##### RBS









|           |   |  |                            |
|-----------|---|--|----------------------------|
| Qy        | 35  | MeIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAsp  | 54                         |
| Db        | 600   | GAGCGCGTGGCAGCGACCGGGTAAAGAAATGGCACCGTGGTCATTTCCGGAAGCCCGG   | 541                        |
| Qy        | 55  | LysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHis | 74                         |
| Db        | 540   | GGCGGGCTGCATTCAGGTGAATATGGAAGAGCTGACCGCGGCTATCATGCTTCAT      | 481                        |
| Qy        | 75  | IleHisGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeu    | 94                         |
| Db        | 480   | GTTACAGAAATCCAAAGCTGCTCGCGGAGAAAAAGACGCAAGATCGTACCGGCTCTT    | 421                        |
| Qy        | 95  | GlnAlaHisGlyHisThrAspProAspLysThrGlyLysHisGluGlyProLeuGlyAsn | 114                        |
| Db        | 420   | GCTGCCCGCGGCATATGATTCGGGTAATACCCATCACCATTAGGCGCTGAAGGTGAT    | 361                        |
| Qy        | 115   | GlyHisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGluThr    | 134                        |
| Db        | 360   | GGACATATGGCGATTTGCCAGCGCTGAGCGCCAATGCTGACGCCAAGGTGAGTGAAACC  | 301                        |
| Qy        | 135   | LeuLeuAlaProArgLeu---ThrValLysGluIleLysGlyArgThrValMetIleHis | 153                        |
| Db        | 300   | GTTTGTCTCTCCATCTCAAGAAATTCGCGGAATCAAGCAGCGCTTCTTGTATGGTCCAT  | 241                        |
| Qy        | 154   | AlaGlyGlyAspAsnTyrSerAspLysProLeuGlyGlyGlyGlyValAlaArgIle    | 173                        |
| Db        | 240   | GTCGAGGGGATATATTATTCGATAGCCTGAGCGGCTTGTCGCGGTGTCGCCGTTT      | 181                        |
| Qy        | 174   | AlaCysGlyValIle-----ProAsn                                   | 180                        |
| Db        | 180   | GCTTGGCGCTGATCGAATAACCGGCCATGCCCAAT                          | 145                        |
| RESULT 21 |   |  |                            |
| AE009694  |   |  |                            |
| LOCUS     | Brucella melitensis 16M chromosome II, section 56 of 107 of the complete sequence.  | 10738 bp   | DNA linear BCT 20-MAR-2003 |
| ACCESSION | AE009694  | AE008918   |                            |
| VERSION   | AE009694.1  | GI:17984749  |                            |
| KEYWORDS  |   |  |                            |
| SOURCE    | Brucella melitensis 16M   |  |                            |
| ORGANISM  | Brucella melitensis 16M   |  |                            |
| REFERENCE |   |  |                            |
| AUTHORS   | DelVecchio, V.G., Kapatral, V., Redkar, R.J., Patra, G., Mujer, C., Loe, T., Ivanova, N., Anderson, I., Bhattacharyya, A., Lykidis, A., Reznik, G., Jablonski, L., Larsen, N., D'Souza, M., Bernal, A., Mazur, M., Goltsman, E., Selkov, E., Elzer, P.H., Hagius, S., O'Callaghan, D., Letesson, J.-J., Haselkorn, R., Kyripides, N. and Overbeek, R. |  |                            |
| TITLE     | The genome sequence of the facultative intracellular pathogen Brucella melitensis   |  |                            |
| JOURNAL   | Proc. Natl. Acad. Sci. U.S.A.   | 99 (1), 443-448  | (2002)                     |
| PUBMED    | 11756688  |  |                            |
| REFERENCE |   |  |                            |
| AUTHORS   | DelVecchio, V.G., Redkar, R.J., Patra, G. and Mujer, C.   |  |                            |
| TITLE     | Submitted (13-NOV-2001) Institute of Molecular Biology and Medicine, University of Scranton, Scranton, PA 18510, USA  |  |                            |
| JOURNAL   |   |  |                            |
| AUTHORS   | Elzer, P.H. and Hagius, S.  |  |                            |
| TITLE     | Submitted (13-NOV-2001) Department of Veterinary Science, LSU Ag Center, 111 Dalrymple Building, Baton Rouge, LA 70803, USA   |  |                            |
| JOURNAL   |   |  |                            |
| REFERENCE |   |  |                            |
| AUTHORS   | Kapatral, V., Loe, T., Ivanova, N., Anderson, I., Bhattacharyya, A., Lykidis, A., Reznik, G., Jablonski, L., Larsen, N., D'Souza, M., Bernal, A., Mazur, M., Goltsman, E., Selkov, E., Haselkorn, R., Kyripides, N. and Overbeek, R.  |  |                            |
| TITLE     | Submitted (13-NOV-2001) Integrated Genomics, Inc., 2201 W. Campbell Park Drive, IL 60612, USA   |  |                            |
| JOURNAL   |   |  |                            |

```

/product="ALKYL HYDROPEROXIDE REDUCTASE C22 PROTEIN"
/protein_id="AAL53819.1"
/db_xref="GI:17984753"
/translation="MLGIGDKLPKPKVGTGKPGFNHBERGVSAFEVTEQSPGKWK
VIFVFKDTEVCPTIAGFARLASEFEDRDAVVLGGSDNFVFLKAWRDHDKLNL
PWSFADTNGSLVDGLVRSPOGATRYTFVDPDNVIOHVYAINLNVGRAPDKTLRV
LUALQDELCPCNREVGGETLKA"
4013..4540
/gene="BMEI10578"
CDS
4013..4540
/gene="BMEI10578"
/codon_start=1
/transl_table=11
/product="AHPD"
/protein_id="AAL53820.1"
/db_xref="GI:17984754"
/translation="NSIDILKSKIPDFAKDVRLNLSMASDETLTPQKYGLFVACGI
ASRNADVRLVAEAKGVDAVIOAKAAASIMGMNVYTFVHLASNDVTFMPAR
LRNVSINPGVDKVPFELWSLVAISNGCMCIDAHEDVLRKANVTAEAIQAAVRPAS
LIOSAAILEAADTE"
complement(4618..5898)
/gene="BMEI10579"
complement(4618..5898)
/gene="BMEI10579"
/codon_start=1
/transl_table=11
/product="TRANSPORTER, MFS superfamily"
/protein_id="AAL53821.1"
/db_xref="GI:17984755"
/translation="MTVAEAVAPQANSVRVLTSAMIGTITTEFFPIYVATRAVIVFP
YVFPASDNGSALLGSFATFAFFARPGVGAIFGHGDKIGRKATLVLAALMTMGLST
VAIGFLVTSIGVAPLALCLQGGLGGEWGAVALLATENAPEGKRTWYGFPP
QUGAPAGFLATGIFLLAETMTEQFAYGWRVPIASAILVAVGLFRLIAETPE
FOKADKARVAVPAQLPRHKNLFLGTIGTMAFVLVFLYMLTFSLQWGLTRALYS
REELVLQIGVIFGLTITPLSALLSDRYGMETIMVIVLIGLIGFIMAPFAAGTA
GVLPFLIGFLMGMTYGPIGAVLAPPTSVRYTGASLAFNLGILGASLAPYIATW
LAIDYGFATVGYIMAAAIISLIGFVFTLKQK"
6727..8211
/gene="BMEI10580"
CDS
6727..8211
/gene="BMEI10580"
/codon_start=1
/transl_table=11
/product="PROBABLY BLUE-COPPER PROTEIN YACK PRECURSOR"
/protein_id="AAL53822.1"
/db_xref="GI:17984756"
/translation="MDAPVLKAPLPLPLVPEASGIIVLKVQKRHSFAKGSTAAS
AGINAGYLGPIVRLMSGESVTLSENAMDEETLLHWGLFVPSHLDGGPHNVIAFGAK
WEPKAVNPFASNPFPHLGHGTARQAHMGVIAGLMIVRDGKDAERGLPETTYGVDDLP
LVLDQRRVLEGDAVAPDMDLIHFGFDWLIANGAIAPEARVPAAMVRLRLINGANA
RNFIHFADRPLLVIASDGGFISQPVSTIEQLTISPERYEVLDVFSNGEAVDLVTYG
DNGSGDGLHMFVTDPALEGKRVKPPVSLDGPAAPEKLSVQRSSFEDERMANMK
LMWROPSSNPHASGDDMDHMGSMAGMDHMGSRADAGPALDALTSGVQMAIAD
KPFDMERIDVEAKLSWETWELTSEMAHPFHIGASFRILSNMGKKPPAHQTGWKDT
ALIDGKAEILVHFDREAKSHPFNFHCHLLEHEDVGMMQAQFVTV"
8481..9005
/gene="BMEI10581"
CDS
8481..9005
/gene="BMEI10581"
/EC_number="1.15.1.1"
/codon_start=1
/transl_table=11
/product="SUPEROXIDE DISMUTASE (CU-ZN)"
/protein_id="AAL53823.1"
/db_xref="GI:17984757"
/translation="MMKSLFIASITVLMVAFAPAFABSTTVVMYEAALPTGPGKEVGVVI
SEAPGLGHFKVNMKLTLPYGHFVHNENPSCGKGVKIVPALAAGHYDFGNTHHH
LGPEGGHGMDDLPRISANADGVSETVWAPHLKLAIEIKORSLMVHVGNDYSDKPEP
LGCGGARFACGVIE"
complement(9487..10590)
/gene="BMEI10582"
complement(9487..10590)
/gene="BMEI10582"

```

```

/EC_number="1.5.3.1"
/codon_start=1
/transl_table=11
/product="SARCOSINE OXIDASE BETA SUBUNIT"
/protein_id="AAL53824.1"
/db_xref="GI:17984758"
/translation="MGPIIDKVTATBESLPQETRIALIGGGVIGVSTALFLAERGIPVT
LFEKEIAGEQSSRNWRCRDSREMPLIVESMRLWEGMNAWGRTGFRVKGIV
YAEKEOMARYAEWIKIARQHGIETDLNGQAEALAPGLTRLAGSSTPLDGRAE
POKAVPMAAKAQEDGAVILQNCVARGIETTRDVSALTEKGRVCEAVVAGGWS
RLIISQKHLPLQKVRSSVFRPTAPTEGGVEFAIAPSDFALRRRDGGVTVASLGSII
ADIVPSFRFRDFTPSLSTEWRSRFLGRFEEAFQWKLRLPQMSIFEKIRTLID
PAPHHTANVQVLAQLKQSRLLRQPSLLRNGP"
Alignment Scores:
Pred. No.: 6,23e-22 Length: 10738
Score: 421.00 Matches: 89
Percent Similarity: 59.90% Conservative: 26
Best Local Similarity: 46.35% Mismatches: 65
Query Match: 44.32% Indels: 12
DB: 1 Gaps: 4
US-10-009-916A-1 (1-180) x AE009694 (1-10738)
QY 1 MetLysIleLysLeuPhe---PheValThrSerIleValThrIleSerLeuLeuThrSer 19
Db 8445 GTTCTCTCGAAGTCTGTCTAACTATATTCAGGAGAGTGAATGAAGTCTTATTATGCA 8504
QY 20 IleThrSerValValLeuAlaCys-----SerValThrSerGluValHis 34
Db 8505 TCGACAAATGCTGTTATGCTTTTCGGGCTTTCGCGAGAAAGACGACGCGTAAATGAT 8564
QY 35 MetIleAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAsp 54
Db 8565 GAGCGCTCCGACCGGACCGGTTAAAGAAATGGCACCGTGTCTATTTCCGAAGCCCG 8624
QY 55 LysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHis 74
Db 8625 GCGGGCTGCATTTCAGGTGAATATGGAAGCTGACCGCGGCTATCATGCTTTCAT 8684
QY 75 IleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeu 94
Db 8685 GTTCAGAAATCCAAAGCTGCTCCGGGAGAAAAAGACGCAAGATCGTACGGCTCTT 8744
QY 95 GlnAlaHisGlyHisThrAspProAspLysThrGlyLysHisGlyGlyProLeuGlyAsn 114
Db 8745 GCTGCGCGGGCATTTATGATCCGGTAAATACCCATCACCATTAGGGCTGAAGTGAT 8804
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
Db 8805 GGACATATGGCGGATTTGCCACGCTGAGCGCAATGCTGACGGCAAGTGTAGTGAACC 8864
QY 135 LeuLeuAlaProArgLeu---ThrValLysGluLysGlyArgThrValMetIleHis 153
Db 8865 GTTGCTGCTCCACATCTCAAGAAATGGCGGAAATCAAGACGCTTCTTTGATGGTCCAT 8924
QY 154 AlaGlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyClyAlaArgIle 173
Db 8925 GTCCGAGGGGATAATTAATTCGATAAGCCTTGAGCGGCTTGTGGCGGTGGTCCCGTTT 8984
QY 174 AlaCysGlyValIle-----ProAsn 180
Db 8985 GCTGCGGGGTGATCGAATAACCGGCCATGCCCAAT 9020

```

RESULT 22  
AF007380  
LOCUS

AF007380 3333 bp DNA linear BCT 30-SEP-1999  
Salmonella typhimurium lambda phage K tail component homolog gene,  
partial cds, lambda phage L tail component homolog, copper-zinc  
superoxide dismutase (sodC), attachment and invasion protein  
homolog and lambda phage M tail component homolog genes, complete  
cds, and lambda phage H tail component homolog gene, partial cds.



```

/note="identified by Glimmer2; putative"
/codon start=1

```

```

/transl_table=11
/product="hypothetical protein"
/protein_id="AAF94728.1"
/db_xref="GI:9656083"
/translation="MGRLSKEDRVKMPPIVSVGLLCSNVFMFTFANYAHLKELNNKP
WITAAVSWGIALPEYLLQVPAANRIGYALTALNVGQKILQEVITLMVFVPSVFLREP
LKIDYLMAGLCVAGAVFAFRSKFA"
/complement(531. .989)
/gene="VC1575"
/complement(531. .989)
/gene="VC1575"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF94729.1"
/db_xref="GI:9656084"
/translation="MSNTHQGLKAVALESAGKIVSLLLGLLHHVAGDSIQLLLLLL
LQHLHNPASGYPEKLLHQAGLTFHFNLMNWAGALVYGAIRLIAYGLMHNLTWM
FALLSGAIYLPPEVYELTFHPGVFSMAALLINLVILMYRIIRSKVPQR"
/complement(1019. .1336)
/gene="VC1576"
/complement(1019. .1336)
/gene="VC1576"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF94730.1"
/db_xref="GI:9656085"
/translation="MSISQITTFEFLTPLVASQKTTIRKSESHPVGPTRVEVFT
LETQKVCIEIDILAVEPLKDFEINEFAEQEAIELPKLALIQIYFNIDELYVITQY
LAK"
/complement(1492. .2313)
/gene="VC1577"
/complement(1492. .2313)
/gene="VC1577"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF94731.1"
/db_xref="GI:9656086"
/translation="WRIFWLKALTSGLFPLSLKNGLCRLVAKPISRKKNAAALTQL
NVALPDLGDVKKQATVSRSLKNLIGLFCGLKRYQVQEPDLVQELIDNQGQGI
VCPMGVYDVYTWLNQGGKAVITIFGAGSSGDRPDENAMISOAKLAGVPLRKON
LMLELAQRIKQGEVWLHDMRTGVPVRFQATQLSATPEFLAHKLACPLYEHYAL
SEGRTQRUHFGRFALHQTDLSRNIAQQAQQLADNMQQAITAHPEQWTLVRRFK"
/complement(2315. .2551)
/gene="VC1578"
/complement(2315. .2551)
/gene="VC1578"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF94732.1"
/db_xref="GI:9656087"
/translation="MKQAIENILIERLQTSIEGSIILTNKFFDFEFSFIDIVAKV
ESQSAQNLFDMPUTMESSVNEVIDMLVSEVGE"
/complement(2535. .4205)
/gene="VC1579"
/complement(2535. .4205)
/gene="VC1579"
/notes="similar to GB:M60177 SP:P11454 GB:M17354 PID:145843
PID:551801; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="enterobactin synthetase component F-related
protein"
/protein_id="AAF94733.1"
/db_xref="GI:9656088"
/translation="MPYNSATDLLQREGYGVVHNFQFINNMAMYCETRHSILAAIQANR

```

```

HAPALWVKQTYTYQEMTDMALSVDYWHLQGVORVAILSVRDLAAYSATWASVLGGM
TYIPLNARATTEQIETLIATQCDSIMVDAQSLSSLETCIDRLHIYALPDVDE
PLRQYPOHFTHTVQITQOVVELLVVXVHLDNEHYAYINQTSGSTKPKRIAYSYN
LHCYISQIDKLPLNAQDRVQSDLTDFDLSHDI FYSLISGACLYVVPBELAKLSPAE
FIHHQTLWMLSVPTVELAQORQTLTSHLSLPSLRFQCGQALLHDLAEQWQATQO
FVINGYPTCTTIAVTHRVFAHSGMASVPIGRAPFECECLAINEQEELMRPEAPEG
YRGELLSGKQLVKGYLNDPLNTQSAFFQHEGRWYSGSDIVTKXSGVLLHGRDHQ
VKIAGRVELEEIEITVVRVTOAHSAIVAPWPLSESYGASGTAVAFVDTHTQWQPDLM
SQCKQNLNPTFVPRWYAIERQLPRNANGKTDIKALQQLASQTYVETSH"
/complement(4269. .5180)
/gene="VC1580"
/complement(4269. .5180)
/gene="VC1580"
/notes="similar to GB:X66866 PID:403331 SP:P52692;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, LysR family"
/protein_id="AAF94734.1"
/db_xref="GI:9656089"
/translation="MLTKRSIKENKAMRHNLHFLHYFWTVAKEGHLTRAAQQLNVS
QSALSSQIRQLEGQLHDLFVROGRSLQTEVGHVLEVAESI FNLSGELLISIMESGE
LQRVOQLRIGSVATLSNFQENFLRPVLGKSDVRLIILSSSTLELLGQLQWMDLIL
SNRPVSDSQPRWCKRIAGQSVCLVCPSPDTLRRYQPPQDLPKVKTLPVCGASSEIRL
QFDLYCEQLGLAVTPYAEVDMMAMRLRLADVEGVTVVPEVVVQDEIETGERLCNYGL
DAVESFYAITTKRHPDMSIVNRLLDN"
/5260. .6810
/gene="VC1581"
/5260. .6810
/gene="VC1581"
/notes="similar to SP:P33607 PID:397909 GB:U00096
PID:1788614 PID:1799638; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase, putative"
/protein_id="AAF94735.1"
/db_xref="GI:9656090"
/translation="MSELSQMLLVSLYGLGLVLSLWLPQQQWRIRAKFATLAGFTAML
GNTLFAIGQASINLPVDGLGIVMAILVSLIGWITRFAHRYLQGEAQARFVTAFLF
TLCVALLVTSOHLVIVLWASGTVSLGRLITFYPERKAARVVAHKFVIRFAEIC
LLALGLIYLDVGSLSLQEIHFALNAQDSPLSLTGAALLAFAILKTAOLPLHGWL
IQWAEPTVSAUHLHAGVVMGGMFGLVLELHLLVAHLSIKAYAFSLSGEAVERA
LTRISIKRVIAWSTCAQMGFMELGLVLELHLLVAHLSIKAYAFSLSGEAVERA
RLHDFQPSQVNGKAVGLAVLSAALIILHLLWQSLTSLNHLPNVAIVAILGFA
PLMWQSGRFPGLSGVLRVALLTOLYFVWMLLFLAGLAPELPASPLALSFVWFCFI
TLXIMQVLLCLYPKSRLSLWLYPWVHGFYLDLFTRLTKYWPVQLTLPALSSQQA
OPTTGERV"
/6735. .9323
/gene="VC1582"
/6735. .9323
/gene="VC1582"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF94736.1"
/db_xref="GI:9656091"
/translation="MARAINSAVRSFINSSATDRRTMNAAPRSLSALSTTANPALST
LVEQVCALIANWPLDRWIVSVYKRIKDPFAAAELKOLAAASPMMTMTLSYHLRW
QNKQIQSADLQQAIAEONSDSESLTIAALQOPTAPSHWPLLCDTVDSRDLHEHPA
WMDAITHAIISOFCAYFDHQADWSPQQTGLFATWREAMIHDSRTITULLNETSVKQI
ATKLPEDAMAIIEQTLAQAPAQEITYLOAVLMRISGWASWCAYLAWAQAGFGRHD
EHLRDLALIRLCWENLLDDGMRGSMVLMQWQSWAPRQSCEDRALRIALLWQRSAE
IAYORQLFAALTSAQESAPQSSYPVEQAAFCIDVRSEVIRRHLEAQPHIQLTGAFG
FGLPIRVOLGCTEASRPLGALLAPSLIVSDTGNEDQDAKLARRLRKRHSWRA
PHLIPASTFTIVETTGAYLTKLKRILSYALSASVEREFAFTEHMQSVKQPTDRP
QTLAQRAQMAANILRGLATEQARLVLLVGHGSQIQNNPORAGLDCGACCGQSGEVN
ARTTAAALINDQAVRQALPEYGISLRDDVHFIAALHNTTTEAMRULFDRHEIIPSHREAR
EQLDQQLTAAHGQARAPSLNHNHQAAPPKDNALSAQQLAQFARLRADHWAOIR
PEWGLTNNAAFTIAPRORSKOAKLDGRVIGLFEQGPDPPEGLQLTIQIMTAPMLVTHWI
NMQYFASVDNDRFSGSKTLLHNVGNGVIGLFEQNGDLRCGLAQSLQSGOGWRHEA
LRLIVVIDAPERIEQVWASREVVEHLVKBWLYLARFADQGIIEIYLQGTWRIQTQPS

```

```
SDSSAR"
complement (9368.. .9883)
/gene="VC1583"
CDS complement (9368.. .9883)
/gene="VC1583"
/note="similar to SP:P00446 PID:L50711; identified by

Alignment Scores:
Pred. No.: 1.43e-21 Length: 12372
Score: 417.00 Matches: 84
Percent Similarity: 60.89% Conservative: 25
Best Local Similarity: 46.93% Mismatches: 66
Query Match: 43.89% Indels: 4
DB: 1 Gaps: 2

US-10-009-916A-1 (1-180) x AE004235 (1-12372)
QY 1 MetLysIleLeuPhePheValThrSerIleValThrIleSerLeuLeuThrSerIle 20
Db 9904 CTTCAATAGGATATCCATCAATCAATCAATTTACTTTGTTGCTGCTGCTTCTTC 9845
QY 21 ThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAspGly 40
Db 9844 TCTTCTCGGTTTGCTCAAGAAATGACTGTG-----GTGAGCGGATTGAGC 9794
QY 41 IleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLys 60
Db 9793 TCTGGCAATCCGTCGACCGTACCTGATAGCGAGTATGTCGCGTCTTTACT 9734
QY 61 ThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlyGlySer 80
Db 9733 CGCGAGTTAACTGCGTTCGACCGCGTTTCACGCGTTTCATGACACGAGATGCTCA 9674
QY 81 CysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyr 100
Db 9673 TGTGAACCTCGTCCAAAGATGCAAAACCGATTAGTGGCGCTCGGTCGGCGATTAT 9614
QY 101 AspProAspLysThrGlyLysHisGlyGlyProLeuGlyAsn---GlyHisLysGlyAsp 119
Db 9613 GATCGCAAAACACGGGAAACACGGCTATCCGTGCAACACGACAAATCACITGGGTGAT 9554
QY 120 LeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArg 139
Db 9553 CTCACGAGCTTTTACGTGATGCGCAAGCCCAAGCAAAATCAACCCGTTTGGCATCCGC 9494
QY 140 LeuThrValLysGluLysGlyArgThrValMetIleHisAlaGlyGlyValIle 178
Db 9493 TTCAAATGCGAGAGTAAAGTAAAGCTTTGATGCTCCACGCGCGGGATACCAT 9434
QY 160 SerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db 9433 TCCGACCATCAATGCTTTGGTGGCGGTGGTGGCGCATCGTATGTTGTTGATT 9377

RESULT 25
AE005367
LOCUS
DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 2 of 3, section 36
of 55.
ACCESSION AE005367
VERSION AE005367.1 GI:12515325
KEYWORDS
SOURCE
ORGANISM Escherichia coli O157:H7 EDL933
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 (bases 1 to 11274)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamouisis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
```

```
JOURNAL Nature 409 (6819), 529-533 (2001)
MEDLINE 21074935
PUBMED 11206551
REFERENCE 2 (bases 1 to 11274)
AUTHORS Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamouisis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2000) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
Location/Qualifiers
1..11274
/organism="Escherichia coli O157:H7 EDL933"
/mol_type="genomic DNA"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:155864"
/note="enterohaemorrhagic"
214..585
/gene="Z2333"
214..585
/gene="Z2333"
/function="orf; Unknown function"
/note="Residues 1 to 118 of 123 are 97.45 pct identical to
residues 1 to 118 of 377 from Escherichia coli K-12 Strain
MG1655: B1377"
/codon_start=1
/transl_table=11
/product="partial putative outer membrane protein"
/protein_id="AAG56384.1"
/db_xref="GI:12515326"
/translacion="MKSQKALLIPALLGAGAAHAAVYNKDNKLDLYKVDGLHYF
SDNSAKDQDSYARLGFKEGTQINDQLTGYGQWEYNIQANNTSSKNQSWTRLAFAGL
KFDSDYGFYGRNYGLDRYAA"
560..1327
/gene="Z2334"
560..1327
/gene="Z2334"
/function="orf; Unknown function"
/note="Residues 2 to 255 of 255 are 98.03 pct identical to
residues 124 to 377 of 377 from Escherichia coli K-12
Strain MG1655: B1377"
/codon_start=1
/transl_table=11
/product="partial putative outer membrane protein"
/protein_id="AAG56384.1"
/db_xref="GI:12515327"
/translacion="MGWTDMLPEFGGDSYTNADNMTGRANGVATYRNTDFGLVNGL
NFAVXZGNNEGASNGQGTNNRDRHENGDWGLSTYDLMGFSAGAXYTSXDR
NDQVHTAAGGKADATAGLKYDANNIYLATMYSETRNTMTPFGSDYAVANKQNF
VTAQYQDFGLRPVPSFLMSKGRDLHAAGAXNPAGVDDKDLVKYADVATYFNKNM
STYVDYKINLLDDEDDSFYAANGISTDDIVALGLVYQF"
1396..1902
/gene="ynaP"
/note="synonym: Z2335"
1396..1902
/gene="ynaP"
/function="orf; Unknown function"
/note="Residues 1 to 168 of 168 are 99.40 pct identical to
residues 1 to 168 of 168 from Escherichia coli K-12 Strain
MG1655: B1376"
/codon_start=1
/transl_table=11
/protein_id="AAG56385.1"
/db_xref="GI:12515328"
/translacion="MNSVITKVSQVTLVADTTGGPMNTILVPIIDISSETQVR
ISHVEAKIDDAEHLTVIPSPYASGLAYSAELPAMDDLKAESQLEIIEKK
FKLPTDRVHVHEEGSPKDRILELAKKIPAHMIIASHRPDITTYLLGSNAAAVVRHA
ECSVLVVR"
2128..>11274
misc_feature
```



```

/notes="Cryptic prophage CP-933R; includes one copy of the
26 bp direct repeat that flanks the prophage; related to
K-12 W31655 prophage Rac"
gene
complement (2483. .3124)
CDS
/genes="Z2337"
/genes="Z2337"
/function="orf; Other or unknown (Phage or Prophage
Related)."
/notes="Residues 80 to 210 of 213 are 41.98 pct identical
to residues 8 to 136 of 140 from GenPept 118 :
gi|4126792|dbj|BAA36750.1| (AB016764) ORF4 [Escherichia
coli]"
/codon_start=1
/transl_table=11
/product="unknown protein encoded by prophage CP-933R"
/protein_id="AAG56386.1"
/db_xref="GI:12515329"
/translation="MPVLTLPYILPGVSFLSDIPQETLSIRNQTIRGEAQIRIGELM
VSRPMQVNGYFMGSLNODGLSDNIQIGLQYIEHRIERTLNHSLTSREVTIVLREIEM
LENMDDLSNVLELELLDKIEVCANVEHAOLQVPSLRTCPTVLCEPEDGVFVRNSMN
SNVCMLYDKALHLVKTAAHPLRSIAVSMIVGRDNCAPDPDRGNFVKN"
complement (3206. .3835)
CDS
/genes="Z2338"
/genes="Z2338"
/function="orf; Other or unknown (Phage or Prophage
Related)."
/notes="Residues 72 to 206 of 209 are 91.11 pct identical
to residues 1 to 135 of 140 from GenPept 118 :
gi|4126792|dbj|BAA36750.1| (AB016764) ORF4 [Escherichia
coli]"
/codon_start=1
/transl_table=11
/product="unknown protein encoded by prophage CP-933R"
/protein_id="AAG56387.1"
/db_xref="GI:12515330"
/translation="MPVTLTSPISOLSPAGVQSLQDAARLESIGIRISGQYSVH
YVOLLDFGSPVPGVLLDRLGLGHRMREAVALEQLNGGVDFLSVNNYFQSVMA
EHRNKYSNKLIMKINSCLRPSPNHFSCPEFLTCPTLDTBTGTGVENSRGAEI
CSLYDKALVOLVETGGHPLSRPITESIMRNKDECHFDYKRAFCCK"
complement (3908. .4483)
CDS
/genes="Z2339"
/genes="Z2339"
/function="orf; Other or unknown (Phage or Prophage
Related)."
/notes="Residues 72 to 187 of 191 are 35.04 pct identical
to residues 20 to 136 of 140 from GenPept 118 :
gi|4126792|dbj|BAA36750.1| (AB016764) ORF4 [Escherichia
coli]"
/codon_start=1
/transl_table=11
/product="unknown protein encoded by prophage CP-933R"
/protein_id="AAG56388.1"
/db_xref="GI:12515331"
/translation="WPLTSDIRSHSNFLGVFVRARIIVANGRGDITVGGTQSVIVYDS
TNGRFSGSGNGGLLSLLLGNSGPRALSMLSDSGEAQSQSIQNKLSOCK
FSVPERLQCPLEALQCPITLEQPEKGI FVKNSDGSVDCTLFDAAFSELVGEGLPHF
LTRPFIASIVKHEECILDDTRGNFVIGN"
complement (4867. .6186)
CDS
/genes="Z2340"
/genes="Z2340"
/function="orf; Other or unknown (Phage or Prophage
Related)."
/notes="Residues 123 to 439 of 439 are 64.26 pct identical
to residues 330 to 645 of 645 from GenPept 118 :
gi|4585436|gb|AAD25464.1|AF125520.59 (AF125520) putative
tail fiber protein [Bacteriophage 933W]"
/codon_start=1
/transl_table=11
/product="putative tail fiber protein encoded by prophage

```

```

CP-933R"
/protein_id="AAG56389.1"
/db_xref="GI:12515332"
/translation="MMMAVKISGVLKDGTPKVENCTIOLKARRXSXTVVVNTVASEN
PDEAGRYSDMVEYQYSVLLVVEGFPFSGHAGTITVVEDSOPGLNDPLGAMSEDDVRP
EALRRFELMVEEARHAEBAKKNVAGEAETSARNAGISASAEBSAANADTSAGDASES
AQOAEASAAAKQSSXSSAAAKQSSXSSAAAKQSSXSSAAAEELSRTAESAGNAARDAT
TATEKAREASQSAEQSRIAAEAENRIPTVVGPPGKEQGPAGPQPKGDKGER
GTGTPVGTATGERPAGDAGPQPKGDKGERGETGLTGNAGPQGPAGTGAAGPAG
PQPKGETGAAGPVGATPGQPKGDGGETQIRFLGPGNIETNSNGWFFPDTDGLALIT
GLTFLAPKATRYQGFQHLQVRFQDGPQWQKGLDVGSDTGRGTE"
complement (6245. .6571)
CDS
/genes="Z2342"
/genes="Z2342"
/function="orf; Other or unknown (Phage or
Prophage Related)."
/notes="Residues 1 to 108 of 108 are 66.66 pct identical to
residues 92 to 199 of 199 from GenPept 118 :
gi|7532789|gb|AAF63231.1|AF151091.2 (AF151091) Lom
[prophage P-Biba]"
/codon_start=1
/transl_table=11
/product="partial putative outer membrane protein Lom
precursor encoded by prophage CP-933R"
/protein_id="AAG56390.1"
/db_xref="GI:12515333"
/translation="XNRNRFWSMAGPSVRVNEWFSAYAMAGVAYSRSVSTFXGDYLRVT
DNKGTCHDLVTGSDDGHRNSTLWAGAGVQFNXPXESVAIDIAVBGSGSGDWRTDGFIV
GVGYKP"
complement (6665. .6946)
CDS
/genes="Z2343"
/genes="Z2343"
/function="orf; Other or unknown (Phage or
Prophage Related)."

```

## Alignment Scores:

| Pred. No.:             | 2,26-21 | Length:       | 11274 |
|------------------------|---------|---------------|-------|
| Score:                 | 414.00  | Matches:      | 88    |
| Percent Similarity:    | 62.50%  | Conservative: | 27    |
| Best Local Similarity: | 47.83%  | Mismatches:   | 53    |
| Query Match:           | 43.58%  | Indels:       | 16    |
| DB:                    | 1       | Gaps:         | 5     |

US-10-009-916A-1 (1-180) x AE005367 (1-11274)

| QY | 1     | MetLysIleLysLeuPhePheValThrSerIleValThrIleSerLeuLeuThrSerIle | 20    |
|----|-------|--|-------|
| DB | 10622 | ATGAAATGTAAATC-----ATTGCTGCATTGCCATCTGCACACA---              | 10663 |
| QY | 21    | ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle        | 36    |
| DB | 10664 | -----GCATCATGCGGATACGCACGACAGAGGAGTCCCAATGAACCTTCTC          | 10711 |
| QY | 37    | AspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspGly    | 56    |
| DB | 10712 | AGTGTGCGAAAAGAGTGCAGCATTTGGAATAAACCATTCAGGACAGACCCCTACCGT    | 10771 |
| QY | 57    | LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis | 76    |
| DB | 10772 | CTGCTGTTACACAGCCCTTCACTCTCTGTCTGAGGACATCATGTTTTCATGTGCAC     | 10831 |
| QY | 77    | GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla | 96    |
| DB | 10832 | GAAAAAGGAATTCGCCGCCGCACTGAAAGACGGAACCCGGTCGACGACATTATCGGT    | 10891 |
| QY | 97    | HisGlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly | 115   |
| DB | 10892 | GGCGGTCTCTTTGACCCGAAAAACACCGGCAAACTCTGGCCCTTGGTCTCCGATGA     | 10951 |
| QY | 116   | HisLysGlyAspLeuProArgLeuValVallysAlaAspGlyIleAlaLysGluThrLeu | 135   |



```

Db      10952  CACCCGGGGACCTCCTCGCGTGTTCGTGACGATGACGAGAAAGCGAACTACCGGTC 11011
QY      136   LeuAlaProArgLeu---ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154
Db      11012  CTTGGCCCGAGACTGAACATCAATAAGAGATTAAGGGCGTCTCTCATGCTTCATGCT 11071
QY      155   GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyGlyAlaArgIleAla 174
Db      11072  GCGCGTGATACCATCATGACCATCCGGAGCCCTGGCGGCTGGTGGTGGCGAATGGCC 11131
QY      175   CysGlyValIle 178
Db      11132  TCGGCATCATTT 11143

RESULT 26
LOCUS   AR204200              34063 bp    DNA    linear    PAT 20-JUN-2002
DEFINITION   Sequence 96 from patent US 6365723.
ACCESSION   AR204200
VERSION     AR204200.1   GI:21500788
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 34063)
AUTHORS    Blattner, F.R., Burland, V., Perna, N.T., Plunkett, G. and Welch, R.
TITLE      Sequences of E. coli O157
JOURNAL    Patent: US 6365723-A 96 02-APR-2002;
FEATURES    source
            1..34063
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.:      6,368-21      Length:      34063
Score:          414.00      Matches:      88
Percent Similarity: 62.50%      Conservative: 27
Best Local Similarity: 47.83%      Mismatches: 53
Query Match:    43.58%      Indels:      16
DB:             6           Gaps:          5

US-10-009-916A-1 (1-180) x AR204200 (1-34063)

QY      1   MetLysIleLysLeuPhePheValThrSerIleValThrIleSerLeuLeuThrSerIle 20
Db      3879  ATGAATGTAAATC-----ATTGTCGCAATTGCCATGCTGCACGCA--- 3920
QY      21   ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle 36
Db      3921  -----GCATCATGCGGATACGACGACGAGAAAGTCCCAATGAACCTTGTC 3968
QY      37   AspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly 56
Db      3969  AGTGCTGACGGAAGAAGTAGTCAGCATTTGAAATAATCAACATTCAGGAGACCCCNACGGT 4028
QY      57   LeuGlnIleLysThrAspLysLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
Db      4029  CTGCTGTTACACACGACCTTCCTCTCTGTCTGAGGCAATTCATGTTTCATGTGCAC 4088
QY      77   GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db      4089  GAAAGGAAATTTGGCCCGGCACTGAAAGACGGAACCGGTCGACGATTCATCGGT 4148
QY      97   HisGlyHisTyrAspProAspLysThrGlyLysHisGlyGlyProLeuGly---AsnGly 115
Db      4149  GCGCGTCACTTTGACCCGAAACACCGGCAACATCTTGGCCCTCGGTCTCCGATGGA 4208
QY      116  HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
Db      4209  CACCCGGGACACTCCCTCGCGTGTGTGACGCGATGACGGAAGAGCAACTACCGGTC 4268
QY      136  LeuAlaProArgLeu---ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154

```

```

Db      4269  CTTGGCCCGAGACTGAACATCAATAAGAGATTAAGGGCGTCTCTCATGCTTCATGCT 4328
QY      155   GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyGlyAlaArgIleAla 174
Db      4329  GCGCGTGATACCATCATGACCATCCGGAGCCCTGGCGGCTGGTGGTGGCGAATGGCC 4388
QY      175   CysGlyValIle 178
Db      4389  TCGGCATCATTT 4400

RESULT 27
LOCUS   BD184769/c              49650 bp    DNA    linear    PAT 17-JUN-2003
DEFINITION   Nucleic acid molecule and polypeptide specific to intestinal hemorrhagic pathogenic Escherichia coli O157:H7, and method of use thereof.
ACCESSION   BD184769
VERSION     BD184769.1   GI:31876969
KEYWORDS    JP 2002355074-A/5.
SOURCE      Escherichia coli O157:H7
ORGANISM    Escherichia coli O157:H7
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1 (bases 1 to 49650)
AUTHORS    Hayashi, H., Shinagawa, H., Makino, K., Hayashi, T., Onishi, S.,
            Hattori, M. and Kurokawa, K.
TITLE      Nucleic acid molecule and polypeptide specific to intestinal hemorrhagic pathogenic Escherichia coli O157:H7, and method of use
JOURNAL    Patent: JP 2002355074-A 5 10-DEC-2002;
            PRESIDENT OF UNIVERSITY OF TSUKUBA
FEATURES    OS Escherichia coli O157:H7
            PN JP 2002355074-A/5
            PD 10-DEC-2002
            PF 24-JAN-2002   JP 2002015959
            PI HIDEO HAYASHI, HIDEO SHINAGAWA, KOZO MAKINO, TETSUYA HAYASHI, SHIN
            ONISHI,
            MASAHIRA HATTORI, KEN KUROKAWA
            PC C12N15/09,C12N15/09,A61K31/7088,A61K39/00,A61K48/00,A61P31/04,
            PC C07K16/12,C12M1/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/
            PC 02,C12Q1/68,
            PC G01N33/15,G01N33/50,G01N33/53,G01N33/53,G01N33/566,G01N37/00,
            PC C12N15/00,
            PC C12N15/00,C12N5/00
            CC Nucleic acid molecule and polypeptide specific to intestinal
            CC hemorrhagic
            CC pathogenic Escherichia coli O157:H7, and method of use thereof
            FH Key Location/Qualifiers
            FT source 1..49650
            FT /organism='Escherichia coli O157:H7'.
            FT Location/Qualifiers
            FT 1..49650
            FT /organism='Escherichia coli O157:H7'
            FT /mol_type='genomic DNA'
            FT /db_xref='taxon:83334'

ORIGIN

Alignment Scores:
Pred. No.:      9,148-21      Length:      49650
Score:          414.00      Matches:      88
Percent Similarity: 62.50%      Conservative: 27
Best Local Similarity: 47.83%      Mismatches: 53
Query Match:    43.58%      Indels:      16
DB:             6           Gaps:          5

US-10-009-916A-1 (1-180) x BD184769 (1-49650)

QY      1   MetLysIleLysLeuPhePheValThrSerIleValThrIleSerLeuLeuThrSerIle 20
Db      42113  ATGAATGTAAATC-----ATTGTCGCAATGCTGCACGCA--- 42072

```

```
Qy 21 ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle 36
Db 42071 -----GCATCATCGGATACGACGAGAACAGGAGTCCCAATGAACCTTGTC 42024

Qy 37 AspAspAsnGlyIleLeysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly 56
Db 42023 AGTGCTGACGGAGAAAGTGCAGCATTCGAAAAATAACCATTCAGGAGAGACCCCTACGGT 41964

Qy 57 LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
Db 41963 CTGCTGTTCACACAGCGCCCTCACTCTCTGTCGAAGGCATTCATGTTTTTCATGTGCAC 41904

Qy 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db 41903 GAAAAAGAAATTCGCGCCCGGCACTGAAGACCGGAAACCGGTTCGACGATTCGCT 41844

Qy 97 HisGlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
Db 41843 GCGCGTCACCTTGACCGGAAAAACACCGGCAACATCTTGGCCCTCGTCTCCGGATGGA 41784

Qy 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrIleu 135
Db 41783 CACCGCGCGACCTCCCTCGCTGTCGATGACGGAAGGAAAGCACTACCCGGTC 41724

Qy 136 LeuAlaProArgLeu---ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154
Db 41723 CTGCGCCCGAGACTGAACCTCATTAAAGAGATTAAAGGGCGTTCTCATGCTTCATGCT 41664

Qy 155 GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174
Db 41663 GCGCGTGATACCATCATGACCATCGGAGCCCTCGGCGGTGGTGGTGCAGAAATGCGC 41604

Qy 175 CysGlyValIle 178
Db 41603 TGGCGCATCATT 41592

RESULT 28
AP002554/c 327773 bp DNA linear BCT 20-MAR-2004
LOCUS
DEFINITION Escherichia coli O157:H7 DNA, complete genome, section 5/20.
ACCESSION AP002554 BA000007
VERSION AP002554.1 GI:13360491
KEYWORDS
SOURCE Escherichia coli O157:H7
ORGANISM Escherichia coli O157:H7
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1
Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S.,
Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T.,
Yanamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T.,
Sasakawa, C. and Shinagawa, H.
Complete nucleotide sequence of the prophage VT2-Sakai carrying the
verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
derived from the Sakai outbreak
Genes Genet. Syst. 74 (5), 227-239 (1999)
2
10198780
10734605
Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M.,
Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and
Hayashi, T.
Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain W61655
Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
3
20557356
11108008
Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S.,
Yutsudo, C.H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T.,
Yanamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and
Shinagawa, H.
```

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Complete nucleotide sequence of the prophage VT1-Sakai carrying the  
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli  
O157:H7 strain derived from the Sakai outbreak  
Gene 258 (1-2), 127-139 (2000)  
20564182  
11111050  
4  
Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K.,  
Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T.,  
Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C.,  
Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and  
Shinagawa, H.  
Complete genome sequence of enterohemorrhagic Escherichia coli  
O157:H7 and genomic comparison with a laboratory strain K-12  
DNA Res. 8 (1), 11-22 (2001)  
21156231  
11258796  
5 (bases 1 to 327773)  
Hattori, M., Ishii, K. and Shiba, T.  
Direct Submission  
Submitted (26-JUN-2000) Masahira Hattori, Kitasato Institute for  
Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamihara,  
Kanagawa 228-8555, Japan  
(E-mail: hattori@genome.ls.kitasato-u.ac.jp,  
URL: http://genome.ls.kitasato-u.ac.jp/, Tel: 81-42-778-8194,  
Fax: 81-42-778-8193)  
genome project

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

This work was done in collaboration with Tetsuya Hayashi, Makoto  
Ohnishi, Keisuke Nakayama (Miyazaki Medical College), Kozo Makino,  
Ken Kurokawa, Katsushi Yokoyama, Masashi Tanaka, Takeshi Honda,  
Teruo Yasunaga, Hideo Shinagawa (Osaka University), Takahiro Murata  
(Shinshu University), Chang-Gyun Han, Eiichi Ohtsubo, Toru Tobe,  
Chihiro Sasakawa (University of Tokyo), Hideto Takami (Japan Marine  
Science and Technology Center), Naotake Ogasawara (Nara Institute  
of Science and Technology), Satoru Kuhara (Kuyshu University), and  
supported by the Research for the Future Program of the Japan  
Society for the Promotion of Science.  
Location/Qualifiers  
1. .327773  
/organism="Escherichia coli O157:H7"  
/mol\_type="genomic DNA"  
/strain="O157:H7"  
/sub\_strain="RIMD 0509952"  
/db\_xref="taxon:83334"  
150. .2258  
/gene="ECs1032"  
150. .2258  
/gene="ECs1032"  
/note="similar to YCBY\_ECOLI gi|1787181 percent identity  
99 in 702 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="putative oxidoreductase"  
/protein\_id="BAB34455.1"  
/db\_xref="GI:13360492"  
/translation="MNSLFASTARGLEBLKTELENLGAECVQVQGVHFKGDTRLV  
YQSLMWSRLASRIMLPLGECKVSDLDLGLGVQAINWTEPNPGATEFAVHFGSLNDTI  
RNSQYGAHKYDAIVDAFTKRLPRNVDRDAPDIRVNVMLHKTASTALDLSGDLGLH  
LRGYRDRAGIAPIKETLAALVMSRQPGTFLDDPMCGSGCTLLIERAMLATDRAPGL  
HRGWGSGWAQDEAIWQEVKBAQTRKRLAEYSGPGLRDLSEPALIALSLGLRIMK  
LAGIGELITTFEVDVAQTPLPKPEYGTVLNSPPYGERLSEPALIALSLGLRIMK  
NQFGWNLISLFSASFDLISCLQLRADKQYKAKNGPLDCVKNVYVAESTPDSKPMVA  
DEANRLRKLKFEKRWAGEIECYRLVDADLEPKYKNVAVDVAWVQVYAPPKTI  
DAKARORLFDIIAATISVLGIAKVLKTRERQKNQYOKLGEKGEFLEVEYNA  
HLWNVLDYLDLGLFLDRIARMLGQMSKGLDFNLFSYTGSAIVHAGLGASRTT  
VMSRTYLENAERLNGLTGRAHRLIQADCLAWLEANEQFLIFDIPPTFSNKR  
MEDAFDVQRDLALMKDLKRLLRAGGTIMFSNNKRGFMDDGLAKLGLKAEITQKT  
LSODFARNRQIHNCWLITAA"  
2270. .4177  
/gene="ECs1033"  
2270. .4177  
/gene="ECs1033"

FEATURES  
source

1. .327773  
/organism="Escherichia coli O157:H7"  
/mol\_type="genomic DNA"  
/strain="O157:H7"  
/sub\_strain="RIMD 0509952"  
/db\_xref="taxon:83334"  
150. .2258  
/gene="ECs1032"  
150. .2258  
/gene="ECs1032"  
/note="similar to YCBY\_ECOLI gi|1787181 percent identity  
99 in 702 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="putative oxidoreductase"  
/protein\_id="BAB34455.1"  
/db\_xref="GI:13360492"  
/translation="MNSLFASTARGLEBLKTELENLGAECVQVQGVHFKGDTRLV  
YQSLMWSRLASRIMLPLGECKVSDLDLGLGVQAINWTEPNPGATEFAVHFGSLNDTI  
RNSQYGAHKYDAIVDAFTKRLPRNVDRDAPDIRVNVMLHKTASTALDLSGDLGLH  
LRGYRDRAGIAPIKETLAALVMSRQPGTFLDDPMCGSGCTLLIERAMLATDRAPGL  
HRGWGSGWAQDEAIWQEVKBAQTRKRLAEYSGPGLRDLSEPALIALSLGLRIMK  
LAGIGELITTFEVDVAQTPLPKPEYGTVLNSPPYGERLSEPALIALSLGLRIMK  
NQFGWNLISLFSASFDLISCLQLRADKQYKAKNGPLDCVKNVYVAESTPDSKPMVA  
DEANRLRKLKFEKRWAGEIECYRLVDADLEPKYKNVAVDVAWVQVYAPPKTI  
DAKARORLFDIIAATISVLGIAKVLKTRERQKNQYOKLGEKGEFLEVEYNA  
HLWNVLDYLDLGLFLDRIARMLGQMSKGLDFNLFSYTGSAIVHAGLGASRTT  
VMSRTYLENAERLNGLTGRAHRLIQADCLAWLEANEQFLIFDIPPTFSNKR  
MEDAFDVQRDLALMKDLKRLLRAGGTIMFSNNKRGFMDDGLAKLGLKAEITQKT  
LSODFARNRQIHNCWLITAA"

gene

150. .2258  
/gene="ECs1032"

CDS

150. .2258  
/gene="ECs1032"

gene

2270. .4177  
/gene="ECs1033"

CDS

2270. .4177  
/gene="ECs1033"

```

/notes="similar to UUP_ECOLI gi|1787182 percent identity 99
in 635 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative ATP-binding component of a transport
system"
/protein_id="BAB34456.1"
/db_xref="GI:13360493"
/translation="MSLISMGAWLSFSDAPLLDNLAEHLIEDNRVCLVGRNGAGKST
LMKILNREGLDDGRIIVQEDLLVARLQDDPPNVEGVSVDVFAEGIERQAEYLKRYH
DISSLVMDPSEKNNELAKVQEQDLHNNMQLENRINEVLAQGLDDENVALSUSGG
MURKALGALVSNRVLDDDEPTNHLDDIETDLEGFKTFTNGTIIIFISHDRSFIRN
MATRVDLDGRKLVTYPNGYDQYLLKEEALRVEELQNAEFKLAQEVIWROGFIKA
RTRNRGRVRLAKWRERGERVGTAKMOVVEASRSGKIVFEMEDVYOVQDKQL
VDFSAQVLRGDKIALIGPGCGKTTLLKMLGQLOQDSGRTHVGTGKLEVAVEDOHR
ELDPKQVMDNLAEQKQEVWNGKPHVLGYLQDLEPHPKRAMTVPRLSGGERNELL
LARLEKSNLULLIDEPNTDNLDELELELLELIDSYQGTVLVLSVDRQFVDNTVTECW
IFEGGKIGRYVGGYHDARGQEQVVALQKQAPVKKNEEPAAPKAETVKRSSKLSYKL
QRELEQLPQLLEDLEAKLEALQTVADASFFSQPHEOTQKVLADMAAAAELEQAFER
WEYLEALKNKG"
4307. .5560
/gene="ECS1034"
/CDs
4307. .5560
/notes="similar to PQIA_ECOLI gi|1787183 percent identity
99 in 417 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="paraquat-inducible protein A"
/protein_id="BAB34457.1"
/db_xref="GI:13360494"
/translation="MCEHHAAKHILCSQCDMLVALPRLEHQKACRCRGTTLVAV
DAPRORPTAYALALFLLLSNLFPVNVNAGVTSEITLLEIPGVLSSEYASLGTP
ELLFVQLVPAYCLTILLVNRALFVRKEQLARVLQKLTWGAELFLAGLVSVF
KLMAYSGVSSFTFVLCVLQAFQCDVRWLDWDIAPEMLRQPLKPGVGTIR
QGLRSCSCCTAILPADEPVCRCGTGYVRRENSLQWTIALIVTSIMLYLPANILPI
VTLDSKMPSTILAGVILLSEGSYPVAAVIFLASIMVPTLKMIAIWLCDWAKGHG
KDRSERMLIYEVVEFVGRWSMIDVFLVIALSVLRMGSLMSIYPAMGALFALVIM
TMFSAMTDPRLSDRQPESEHEES"
5565. .7205
/gene="ECS1035"
/CDs
5565. .7205
/notes="similar to PQIB_ECOLI gi|1787184 percent identity
99 in 546 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="paraquat-inducible protein B"
/protein_id="BAB34458.1"
/db_xref="GI:13360495"
/translation="MESNNGEAKIQKVNWSVPVIFPIVITALIGAWLVIFYHSHQGE
PVLITANAEIGEGKTTIKRSVDVGVESATLADDLTHVEIKARLNSGMEKLLKHD
VFWVVPQIGEGSLTSLGSVYIELPGAKGSKMDYLLDPLPAPPAKIRV
VLDSKKAQLSGDPVLPFRGVSVETSTFDQKRNSYOLFVNPYDLVLTNSVR
WKDSGIADVLTSAGMRVEMSLTTLTSGSVSDPVEGLDLPQVAPKTAFLVYDDKS
IODSLYTHIDYLMFKDSVRLGQAPVEFRGLRGTVSKVPFAPNROTDFNDYR
IPVLTIRIFERLKMQLGENADVVEHLGELLKRGSLKGNLVGALYVLDLFFPNT
FALTIRIFENGQIIPTVSGGLAQIQRLMEALDKINKPLNPMIEQATSLSESQRT
MKNLQTTLLDSNKNILASQMQLPDMDQSTRELNRSMQGFQPGSAAYKNKVMADQRL
DQVRELQPVKLTKNEKSNALVFEAKKDKPEPKAKQ"
7217. .7765
/gene="ECS1036"
/CDs
7217. .7765
/notes="similar to YMBE_ECOLI gi|1787185 percent identity
98 in 182 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"

```

```

/protein_id="BAB34459.1"
/db_xref="GI:13360496"
/translation="MTIALWLAGCSEINKYNYQLPVVQSGSTOSTASQGNLLWVE
QVAVPDYLAGNVVYQTSQDVKYVIANNIMWSPQLQQLANTIVANLSTQIPGWVASQ
PLGSAQDTINVTIEFNKRYDGKIVSGEWLNLHQQLIKRFRFLGVGTQDGYDEM
KVLAVQWSQEAASIAQEIKRUP"
8021. .8188
/gene="ECS1037"
/CDs
8021. .8188
/notes="identical to RMF_ECOLI gi|1787186 (Conserved in
E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/product="ribosome modulation factor"
/protein_id="BAB34460.1"
/db_xref="GI:13360497"
/translation="MKRQKRDRLERAHQRYQAGIAGRSKEMCPYQTLNQRSQWLGG
US-10-009-916A-1 (1-180) x AP002554 (1-327773)
Alignment Scores:
Pred. No.: 5,628-20 Length: 327773
Score: 414.00 Matches: 88
Percent Similarity: 62.50% Conservatives: 27
Best Local Similarity: 47.83% Mismatches: 53
Query Match: 43.58% Indels: 16
DB: 1 Gaps: 5
QY 1 MetLysIleLysLeuPheValThrSerIleValThrIleSerLeuLeuThrSerIle 20
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 66038 ATGAATCTGAATC-----ATGTGTCATTCATTCATTCATTCATTCATTCAT 65997
QY 21 ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle 36
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 65996 -----GCATTCGCGATACGACGACGACGACGACGACGACGACGACGACGAC 65949
QY 37 AspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspGly 56
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 65948 AGTGTCTGACGAAAGAAAGTTCGACATTCGAAATAAACATTCAGGACGACCCCTACGGT 65889
QY 57 LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 65888 CTGCTGTTTACACACGACCTTCACCTCTCTCTGTCTGAGGACATTCATGTTTTCATGTGAC 65829
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 65828 GAAAAAGGAAATTTGCGCCCGCGCACTGAAAGACGAGAAACCGGTGCGACGACATTCATCGGCT 65769
QY 97 HisGlyHisTyrAspProAspLysThrGlyLysHisGlyGlyProLeuGly---AsnGly 115
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 65768 GCGGTCTACTTGTACCCGCAAAACACCGCGCAACATCTTGGCCCTCTGCTCCGATGGA 65709
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 65708 CACCGCGCGACCTCCCTCGCTGTTCTGTCGACGATGACGAGAAAGCACTACCCGCTC 65649
QY 136 LeuAlaProArgLeu---ThrValLysGluLysGlyArgThrValMetIleHisAla 154
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 65648 CTGCGCCGACGACGACATCAATAAGAGATTAAGGCGCTTCTCTCATGCTTCATGCT 65589
QY 155 GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 65588 GCGGTGATAACCATCATGACCATCCGAGCCCTGGCGCGTGGTGGTGCAGAAATGCGC 65529
QY 175 CysGlyValIle 178
||||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 65528 TCGCGCATCATTT 65517
RESULT 29
AF254764
LOCUS
AF254764 718 bp DNA linear PHG 30-AUG-2001

```

```

DEFINITION Bacteriophage Fels-1 putative Cu/Zn superoxide dismutase precursor
            (sodCIII) gene, complete cds.
ACCESSION AF254764
VERSION AF254764.1 GI:9049794
KEYWORDS Phage Fels-1
SOURCE Phage Fels-1
ORGANISM Viruses.
REFERENCE 1 (bases 1 to 718)
          Figueroa-Bossi,N., Uzau,S., Maloriot,D. and Bossi,L.
          Variable assortment of prophages provides a transferable repertoire
          of pathogenic determinants in Salmonella
          Mol. Microbiol. 39 (2), 260-271 (2001)
JOURNAL 21065122
MEDLINE 11136448
PUBMED
REFERENCE 2 (bases 1 to 718)
          Figueroa-Bossi,N. and Bossi,L.
          Direct Submission
          Submitted (10-APR-2000) Centre de Genetique Moleculaire, CNRS,
          Gif-Sur-Yvette 91198, France
JOURNAL
FEATURES
     source
     1..718
     /organism="Phage Fels-1"
     /varion
     /mol_type="genomic DNA"
     /db_xref="taxon:128975"
     /note="found in lysogenic state in Salmonella enterica
     serovar Typhimurium strain LT2 at map position 20 cs"
     152..676
     /gene="sodCIII"
     152..676
     /gene="sodCIII"
     /note="sodCIII"
     /codon_start=1
     /transl_table=11
     /product="putative Cu/Zn superoxide dismutase precursor"
     /protein_id="AAF62484.1"
     /db_xref="GI:9049795"
     /translation="MNKIMTLAVLFSSAASASTDATVNLVNAVGTGOKIGNITIT
     ETEYGLIFTHLSLPAGIHGFHENGSCDAGMKDGKPVAAALAGSLDQPHNKL
     GPYNPEGHLDPLALYVQAGNADYPVLAFLPKSISQVKHALMIHTGGNHDHPNP
     LGGGGARIACGVK"
ORIGIN
Alignment Scores:
Pred. No.: 1,69e-22 Length: 718
Score: 413.50 Matches: 84
Percent Similarity: 61.76% Conservative: 21
Best Local Similarity: 49.41% Mismatches: 62
Query Match: 43.53% Indels: 3
DB: 7 Gaps: 3

US-10-009-916A-1 (1-180) x AF254764 (1-718)
Qy 11 IleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerValThr 30
Db 164 TTAATGACACTGCTGATTTTAA--TTCAAGTAGTCAGCTCAGCAGCAAGTACCGAC 220
Qy 31 SerGluValHisMetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThrPhe 50
Db 221 GCAACGGTTAACTTGTAAATGCCAATGCCACAGGTCAAAAGATCGGCAATATTACAATT 280
Qy 51 ThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGlu 70
Db 281 ACCGAAACAGAAATATGGCTTGTATTATCCACCACATTTGTCTCACTCCAGCCGGTATT 340
Qy 71 HisGlyPheHisIleHisGluGlySerCysGlyProAlaGluHisAspGlyHisLeu 90
Db 341 CATGGGTTTATATCCACGAGAACCGAAGCTGTGATGTCGATGAAGACGGTAAACCT 400
Qy 91 ThrAlaGlyLeuGlnAlaHisGlyHisIleTyraAspProAspLysThrGlyHisGlyGly 110
Db 401 GTCGAGCACTGTGTCGGGAGGGCACCCTTGACCCCTCAGCACACAATAAATCATCTCGGT 460

```

```

Qy 111 ProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129
Db 461 CCGTATAATCCTGAAGGTCACTCGGTGATCTTCGGCGCTGTATGTAAATCAGGCAGGT 520
Qy 130 IleAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLysGluIleLysGlyArg 148
Db 521 AACGCTGATTATCCGGTTCTTGGCCGCCCAATCAATCAATCAATCAATCAATCAATCAAT 580
Qy 149 ThrValMetIleHisAlaGlyGlyAspAsnTySerAspLysProLeuProLeuGlyGly 168
Db 581 GCCATCATGATTCATACCGCGGTGACACCAATGAAGATCATCCCAACCCGCTTGGGGGC 640
Qy 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
Db 641 GTTGGTCAGCAATCGCTCTGGGGTTATA 670

RESULT 30
AE008737/c
LOCUS Salmonella typhimurium LT2, section 45 of 220 of the complete
DEFINITION genome.
ACCESSION AE008737 AE008738 AE008739
VERSION AE008737.2 GI:22024638
KEYWORDS Salmonella typhimurium LT2
SOURCE Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
          Enterobacteriaceae; Salmonella.
ORGANISM 1 (bases 1 to 65219)
          McClelland,M., Sanderson,K.E., Spieth,J., Clifton,S.W.,
          Latreille,P., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F.,
          Hou,S., Layman,D., Leonard,S., Nguyen,C., Scott,K., Holmes,A.,
          Grewal,N., Mulvaney,E., Ryan,E., Sun,H., Florea,L., Miller,W.,
          Stoneking,T., Khan,M., Waterston,R. and Wilson,R.K.
          Complete genome sequence of Salmonella enterica serovar Typhimurium
          LT2
          Nature 413 (6858), 852-856 (2001)
JOURNAL 21534948
MEDLINE 11677609
PUBMED
REFERENCE 2 (bases 1 to 65219)
          The Salmonella typhimurium Genome Sequencing Project
          Direct Submission
          Submitted (29-MAR-2001) Genome Sequencing Center, Department of
          Genetics, Washington University School of Medicine, 4444 Forest
          Park Boulevard, St. Louis, MO 63108, USA
          On or before Jul 31, 2002 this sequence version replaced
          GI:16419409, GI:16419430, GI:16419386.
          COMMENT Supported by NIH grant 5U 01 AI43283
          Coding sequences below are predicted from manually evaluated
          computer analysis, using similarity information and the programs;
          GLIMMER; http://www.tigr.org/softlab/glimmer/glimmer.html and
          GeneMark; http://opal.biology.gatech.edu/GeneMark/
          EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
          Encyclopedia of Genes and Genomes; http://www.genome.ad.jp/kegg/,
          and Pedro Romero and Peter Karp at EcoCyc;
          http://ecocyc.PangeaSystems.com/ecocyc/
          The analyses of ribosome binding sites and promoter binding sites
          were kindly provided by Heladia Salgado, Julio Collado-Vides and
          ReguonDB;
          http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset
          This sequence was finished as follows unless otherwise noted: all
          regions were double stranded, sequenced with an alternate
          chemistry or covered by high quality data (i.e., phred quality >=
          30); an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by sequence
          from more than one ml3 subclone.
          Location/Qualifiers

```

FEATURES

```
source
join(1..13108,54773..65219)
/organism="Salmonella typhimurium Lt2"
/mol_type="genomic DNA"
/strain="LT2; SGSC 1412; ATCC 700720"
/db_xref="ATCC:700720"
/db_xref="taxon:99287"
/focus
/note="LT2"
13109..54772
/organism="Phage Fels-1"
/mol_type="genomic DNA"
/db_xref="taxon:128975"
68..1215
/genes="potG"
/notes="synonym: STM0878"
68..73
/genes="potG"
/notes="putative RBS for potG; RegulonDB:STMSIH001174"
82..1215
/genes="potG"
/note="ABC superfamily (atp_bind); similar to E. coli
ATP-binding component of putrescine transport system
(AAC73942.1); Blastp hit to AAC73942.1 (404 aa), 93%
identity in aa 28 - 404"
/codon_start=1
/transl_table=11
/product="putrescine transporter"
/protein_id="AAL19814.1"
/db_xref="GI:16419387"
/translation="MNDVPHRQVQKTRKALTEPILLENLTKSFDGQHVDDVSLTIYK
GEIFALLGASGKTLRLMLAGFEPPSSQIMLDGVDLAHPVPPYQRPINMMFQSYAL
FPHMTVEQNIAGLKQDKLPKAEITSRVNEMGLVHMQBFAKKPKQLSGGQRQVAL
ARSLAKPKLLDEEPALDKKLRDMQLEVVILRERGVTCVMVTHDQEEAMTMAG
RIAINNRGFVQIPEETIYERPTTYSAEFTGVSNNVFGLLKAREEDGLVIDAPGLV
HLKVDADASVDNNPVVVALRPEKIMLCESPADGYNFAVGEVAHIALVIGLSIYHV
RLKSGMISAQIQNAHRYRKGKLPDWGDEVLCWEADSCVVLIV"
1214..2178
/genes="potH"
/notes="synonym: STM0879"
1214..1219
/genes="potH"
/notes="putative RBS for potH; RegulonDB:STMSIH001175"
1225..2178
/genes="potH"
/note="ABC superfamily (membrane); similar to E. coli
putrescine transport protein; permease (AAC73943.1);
Blastp hit to AAC73943.1 (317 aa), 92% identity in aa 1 -
317"
/codon_start=1
/transl_table=11
/product="putrescine transporter"
/protein_id="AAL19815.1"
/db_xref="GI:16419388"
/translation="MSTLEPPARIKSPGGAALWLRMQMKGRKLVIALPYLWLILFL
LUPFLIVKISAEWARAIPPTDLEWADGQSLITNLGNLQLTDDPEAYELQS
LQIAGITICLLGLYPLAWAHSKPSRNLILLIPLPSWTSFLIRYAWMGILKN
NGALNNLWLVQIPLHTNLNAVYIGIVAYLPEMVLPIYALTALRIDYSLVEAS
LDLGLARPLTFISIIPLTKGGIAGSLMVFIPAVGEFVPELLGGPDSIMIGRLVWQ
EFENNRDPEVASAVALIMLLLLIPLPIMWFHKQKSVGEHG"
2161..3020
/genes="potI"
/notes="synonym: STM0880"
2161..2166
/genes="potI"
/notes="putative RBS for potI; RegulonDB:STMSIH001176"
2175..3020
/genes="potI"
/note="ABC superfamily (membrane); similar to E. coli
putrescine transport protein; permease (AAC73944.1);
Blastp hit to AAC73944.1 (281 aa), 94% identity in aa 1 -
281"
/codon_start=1
/transl_table=11

gene
/product="putrescine transporter"
/protein_id="AAL19816.1"
/db_xref="GI:16419389"
/translation="MNNLFPVSPRILLILVGTFLYAPMLMLVIYSFNSSKLVTVM
AGWSTRYGELFRDTPAMSAVGLSUTIAACAATMAVLGTIAAVVMVRFGRFGANGF
AFMITAPLVDPVITGLSLLFLVALGHAIGWFSRDMGLTIIAHVTFCTAYAVVIA
SRRELDRSIEEAADGAPLKVFVFTLPMIPAVISGWLIAFLTDLDDLVASPV
SGRGATTLPMLVFSSVRMGVNPINALATLILGVGVIGVFIAWYLARAERKQIRDIQ
BARRG"
3079..3567
/genes="ybJ"
/notes="synonym: STM0881"
3079..3084
/genes="ybJ"
/notes="putative RBS for ybJ; RegulonDB:STMSIH001177"
3094..3567
/genes="ybJ"
/notes="similar to E. coli orf, hypothetical protein
(AAC73945.1); Blastp hit to AAC73945.1 (162 aa), 80%
identity in aa 6 - 162"
/codon_start=1
/transl_table=11
/product="putative inner membrane protein"
/protein_id="AAL19817.1"
/db_xref="GI:16419390"
/translation="MGFIKQSTSSHARLNVPVLVOVAAAILIRGLDVLIMINTLGI
RGMGRFIHRSVOTWSLTIVFLASLVFVEIYCAFLVKGRSWRWWVLATOLIISGY
LWAASLGVGPPELFSIAGESKRDILHSLVMQKLPDLLILFLFIPAPSRFRFLQ"
3597..4740
/genes="ybJf"
/notes="synonym: STM0882"
3597..3602
/genes="ybJf"
/notes="putative RBS for ybJf; RegulonDB:STMSIH001178"
3610..4740
/genes="ybJf"
/EC_number="2.1.1.1"
/notes="similar to E. coli putative enzyme (AAC73946.1);
Blastp hit to AAC73946.1 (375 aa), 87% identity in aa 1 -
372"
/codon_start=1
/transl_table=11
/product="putative tRNA (uracil-5)-methyltransferase"
/protein_id="AAL19818.1"
/db_xref="GI:16419391"
/translation="MQCALYDAGRCRSOWITQSVNEQLSAKTADLHRLLAGLPVEQW
CAPIGGQHFPRNKAMVSVSGVERPLFGLMHRDGTVPDLGCGPIYPASDFVFSALK
PFIRAGLTFFVNAKRGELKYLITLTFQFDGMMMLRFLRSETKLTQIRAAALPWLRA
QLPOLRVITANIQPVHMAIMEGETEYILTQQALAEFNDVPLWIRPQSFFQTNPTVA
SRLVATARDWVQQLPVRHMDLFCGVGGFGLHCATPQMLTGIETAPAIACAKQSA
ELGLTRLHFOALDSTQFATQGETPDVLVNPFRGIGKPLCDLYLAQMAPRFIIYSSC
NAQTWAQDIRHLPNYRIQRVQLFDMPPHTAHVEVLALLRSI"
5019..6366
/genes="STM0884"
5019..5024
/genes="STM0884"
/notes="putative RBS for STM0884; RegulonDB:STMSIH001179"
5023..6366
/genes="STM0884"
/notes="similar to E. coli orf, hypothetical protein
(AAC77150.1); Blastp hit to AAC77150.1 (484 aa), 38%
identity in aa 20 - 441"
/codon_start=1
/transl_table=11
/product="putative inner membrane protein; homology to
Sgar from Vibrio"
/protein_id="AAL19819.1"
/db_xref="GI:16419392"
/translation="MEGVOTMFAKPIDVIOITFLTEPAILIGLVGVYALDKKTIPI
ITGMISAMGLMMVLFGGQFSATFKPVAEASVKAIVGHVYLMDSYAKAAATQIALGD
NFGVGVYVLVLAFFTNLLLVFGRTYGAIGFILTGNVSHSQVAILVILVFWLFGWV
OSIVIAGLTGVAFNFTLIVKPIAKVTNNAGFTIAHNMGLWFFSKFAHKFGDPE
KHDAENLKLPGWLAIFNHNVTAIAVMTLVFGGFLLATGIDNVQLMAKGPWYIYIN
```







|   |  |  |                                       |
|---|--|--|---------------------------------------|
| Qy  | 155  | GlyGlyAspAsnTyrSerAspLeuProLeuGlyGlyGlyGlyValaAargileala                   | 174                                   |
| Db  | 7305   | GGCGGTGNTACCATCATGACCATCCGGAGCCCTGGCGGTGGTGGCGAGATGSCC                     | 7364                                  |
| Qy  | 175  | CysGlyValIle 178   |                                       |
| Db  | 7365   | TGCGGCATCAT 7376   |                                       |
| RESULT 22                                     |  |  |                                       |
| BD184776/c                                    |  |  |                                       |
| LOCUS   | BD184776   | 39824 bp   | linear                                |
| DEFINITION                                    | Nucleic acid molecule and polypeptide specific to intestinal hemorrhagic pathogenic Escherichia coli O157:H7, and method of use thereof. |  |                                       |
| ACCESSION                                     | BD184776   |  |                                       |
| VERSION                                       | BD184776.1   | GI:31876976  |                                       |
| KEYWORDS                                      | JP 2002355074-A/12.  |  |                                       |
| SOURCE  | Escherichia coli O157:H7   |  |                                       |
| ORGANISM                                      | Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.                                       |  |                                       |
| REFERENCE                                     | 1 (bases 1 to 39824)   |  |                                       |
| AUTHORS                                       | Hayashi,H., Shingawa,H., Makino,K., Hayashi,T., Onishi,S., Hattori,M. and Kurokawa,K.  |  |                                       |
| TITLE   | Nucleic acid molecule and polypeptide specific to intestinal hemorrhagic pathogenic Escherichia coli O157:H7, and method of use          |  |                                       |
| JOURNAL                                       | Patent: JP 2002355074-A 12 10-DEC-2002;  |  |                                       |
| COMMENT                                       | PRESIDENT OF UNIVERSITY OF TSUKUBA   |  |                                       |
|   | OS   | Escherichia coli O157:H7   |                                       |
|   | FN   | JP 2002355074-A/12   |                                       |
|   | PD   | 10-DEC-2002  |                                       |
|   | PF   | 24-JAN-2002  | JP 2002015959                         |
|   | PI   | HIDEO HAYASHI, HIDEO SHINAGAWA, KOZO MAKINO, TETSUYA HAYASHI, SHIN ONISHI, |                                       |
|   | PI   | MASAHIRA HATTORI, KEN KUROKAWA   |                                       |
|   | PC   | C12N15/09, C12N15/09, A61K31/7088, A61K39/00, A61K48/00, A61P31/04,        |                                       |
|   | PC   | C07K14/245,  |                                       |
|   | PC   | C07K16/12, C12M1/00, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/       |                                       |
|   | PC   | 02, C12Q1/68,  |                                       |
|   | PC   | G01N33/15, G01N33/50, G01N33/53, G01N33/53, G01N33/566, G01N37/00,         |                                       |
|   | PC   | C12N15/00,   |                                       |
|   | PC   | C12N15/00, C12N5/00  |                                       |
|   | CC   | Nucleic acid molecule and polypeptide specific to intestinal               |                                       |
|   | CC   | Nucleic acid molecule and polypeptide specific to intestinal               |                                       |
|   | CC   | hemorrhagic  |                                       |
|   | CC   | pathogenic Escherichia coli O157:H7, and method of use thereof             |                                       |
|   | FH   | Key  | Location/Qualifiers                   |
|   | FT   | source   | 1. 39824                              |
|   | FT   | source   | /organism='Escherichia coli O157:H7', |
| FEATURES                                      | Location/Qualifiers  |  |                                       |
| source  | 1. 39824   |  |                                       |
|   | /organism='Escherichia coli O157:H7'   |  |                                       |
|   | /mol_type='genomic DNA'  |  |                                       |
|   | /db_xref='taxon:83334'   |  |                                       |
| ORIGIN  |  |  |                                       |
| Alignment Scores:                             |  |  |                                       |
| Pred. No.:                                    | 8.79e-21   | Length:  | 39824                                 |
| Score:  | 413.00   | Matches:   | 88                                    |
| Percent Similarity:                           | 62.50%   | Conservative:  | 27                                    |
| Best Local Similarity:                        | 47.83%   | Mismatches:  | 53                                    |
| Query Match:                                  | 43.47%   | Indels:  | 16                                    |
| DB:   | 6  | Gaps:  | 5                                     |
| US-10-009-916A-1 (1-180) x BD184776 (1-39824) |  |  |                                       |
| Qy  | 1  | MetLysIleLysLeuPhePheValThrSerIleValThrIleSerLeuLeuThrSerIle               | 20                                    |
| Db  | 31584  | ATGAATGTAATC-----ATTGTCGCATTTGCCATGCTGCAGCA---                             | 31543                                 |
| Qy  | 21   | ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle                      | 36                                    |
| Db  | 31542  | -----GCATCATCGGATACGCGACGACGAGAGTCCCAATGACCTTGTC                           | 31495                                 |



```

Db      6902 AGTGCTGACGGAAGAAAGTACGATTTGGAATAATACCATTCAGGAGACCCCTACGGT 6961
QY      57  LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
Db      6962 CTGCTGTTTACACACGCCCTTCACTCTCTGCTGAAGGCAATTCATGTTTCAITGACAC 7021
QY      77  GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db      7022 GAAAAGGAAATTTGCCCGCCGACTGAAAGACGGAACCGGTCCGACGATTTATCGGCT 7081
QY      97  HisGlyHisTyrAspProAspLysThrGlyLysHisGlyGlyProLeuGly---AsnGly 115
Db      7082 GCGGGTCATCTTTACCGGAAAAACACCGGCAACATCTTTGGCCCTGGTCTCCGGATGGA 7141
QY      116 HisLysGlyAspLeuProArgLeuValLysAlaAspGlyValLeuAlaLysGluThrLeu 135
Db      7142 CACCTGGGGACCTCCTCGCTGTTCTGTCAGCATGACGGAACGAACTACCCGGTC 7201
QY      136 LeuAlaProArgLeu---ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154
Db      7202 CTGGCCCGGAGACTGAACCTAATAAGAGATTAAAGGCGTCTCTCATGCTTCATGCT 7261
QY      155 GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174
Db      7262 GCGGGTGATAACCATCATGACCATCCGAGCCCTCGGCGGTGGTGGTGCGAGATGGCC 7321
QY      175 CysGlyValIle 178
Db      7322 TCGGGCATCATT 7333

RESULT 34
AP002556/c
LOCUS      291136 bp      DNA      linear      BCT 20-MAR-2004
DEFINITION Escherichia coli O157:H7 DNA, complete genome, section 7/20.
ACCESSION AP002556 BA000007
VERSION    AP002556.1 GI:13361156
KEYWORDS
SOURCE      Escherichia coli O157:H7
ORGANISM    Escherichia coli O157:H7
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.

REFERENCE
AUTHORS      Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S.,
            Kurokawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T.,
            Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T.,
            Sasakawa,C. and Shinagawa,H.
TITLE        Complete nucleotide sequence of the prophage VT2-Sakai carrying the
            verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
            derived from the Sakai outbreak
JOURNAL      Genes Genet. Syst. 74 (5), 227-239 (1999)
MEDLINE      20198780
PUBMED      10734605
REFERENCE
AUTHORS      Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M.,
            Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and
            Hayashi,T.
TITLE        Comparative analysis of the whole set of rRNA operons between an
            enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
            Escherichia coli K-12 strain MGI655
JOURNAL      Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
MEDLINE      20557356
PUBMED      11108008
REFERENCE
AUTHORS      Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S.,
            Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T.,
            Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and
            Shinagawa,H.
TITLE        Complete nucleotide sequence of the prophage VT1-Sakai carrying the
            Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
            O157:H7 strain derived from the Sakai outbreak
JOURNAL      Gene 258 (1-2), 127-139 (2000)
MEDLINE      20564182
PUBMED      11111050

```

```

REFERENCE
AUTHORS      Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K.,
            Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T.,
            Tanaka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C.,
            Ogasawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and
            Shinagawa,H.
TITLE        Complete genome sequence of enterohemorrhagic Escherichia coli
            O157:H7 and genomic comparison with a laboratory strain K-12
JOURNAL      DNA Res. 8 (1), 11-22 (2001)
MEDLINE      11258796
PUBMED      21156231
REFERENCE
AUTHORS      Hattori,M., Ishii,K. and Shiba,T.
TITLE        Direct Submission
JOURNAL      Life Sciences, Kitasato University; Kitasato 1-15-1, Sagami-hara,
            Kanagawa 228-8555, Japan
            (E-mail:hattori@genome.ls.kitasato-u.ac.jp,
            URL:http://genome.ls.kitasato-u.ac.jp/, Tel:81-42-778-8194,
            Fax:81-42-778-8193)
COMMENT      This work was done in collaboration with Tetsuya Hayashi, Makoto
            Ohnishi, Keisuke Nakayama (Miyazaki Medical College), Kozo Makino,
            Ken Kurokawa, Katsushi Yokoyama, Masashi Tanaka, Takeshi Honda,
            Teruo Yasunaga, Hideo Shinagawa (Osaka University), Takahiro Murata
            (Shinshu University), Chang-Gyun Han, Eiichi Ohtsubo, Toru Tobe,
            Chihiro Sasakawa (University of Tokyo), Hideto Takami (Japan Marine
            Science and Technology Center), Naotake Ogasawara (Nara Institute
            of Science and Technology), Satoru Kuhara (Kyuushu University), and
            supported by the Research for the Future Program of the Japan
            Society for the Promotion of Science.
FEATURES
            Location/Qualifiers
            1..291136
            /organism="Escherichia coli O157:H7"
            /mol_type="genomic DNA"
            /strain="O157:H7"
            /sub_strain="RIMD 050952"
            /db_xref="taxon:83334"
            126..380
            /gene="ECs1692"
            126..380
            /gene="ECs1692"
            /note="similar to YMG_ECOLI gi|1787445 percent identity
            97 in 84 aa (Conserved in E.coli K-12)"
            /codon_start=1
            /evidence=not experimental
            /transl_table=11
            /product="hypothetical protein"
            /protein_id="BAB35115.1"
            /db_xref="GI:13361157"
            /translation="MGIIAWIIFGLIAGIILKIMPRDGGGFFLTCLILGIVGAVVGG
            WLATMFGIGSISGFNLHSFLVAVVGVGAILVLGVFLLRRE"
            complement(430..2400)
            /gene="ECs1693"
            complement(430..2400)
            /gene="ECs1693"
            /note="probable TonB dependent outer membrane
            receptor, similar to TonB dependent outer membrane receptor
            PrtA [Escherichia coli CFT073] gi|3661477|gb|AAC61709.1|
            percent identity 97 in 656 aa"
            /codon_start=1
            /evidence=not experimental
            /transl_table=11
            /product="putative TonB dependent outer membrane receptor"
            /protein_id="BAB35116.1"
            /db_xref="GI:13361158"
            /translation="WRLKKHVLCTALSLAFTQAAVAAQESDLSLTWSSPVSSTTTTTL
            DQPTMKALDKONVAQLSVPGVVLQKSSREOVKVRGPDSPVYEDGVPYIVP
            YDGNLDLRLTNLNGAVEKGYSSLLQGNQMGAINITTKTKPLEASLGVRQG
            WSRQNDYDMHSPASSELGYLOVGSQKLQDFLPGHGVNNDIAGHKWIMNSA
            DDKRGIVKLGFTFRENDEYTLTYIKQDGKONPPYSGNSGQKRSYQWPEYDKSFY
            QGTLQNDRLTKSLRYLTFENTLMYNSLADLNKKGSYSHYGDYSDGLQIAD
            VRENDLLTFAVNWKDDVHREKGAHPAAVDREDTWLSASEYQMAADNDVDVAGISY

```

gene  
CDS  
DNRDSVBAKHEKDGSIHYDDNNQAFNMQVMGKYHFVNEIDTLLSYIDRKRPPTLK  
ERYTTSPAYNOIAIVPOLKPERARGVDLTWNGAFTHDWGFVSVYVNRVSDAILGH  
NIDATTQNGSVGVSGLDAGIKGKISLIDVGLSVLHADAKRKDGKIDLP  
QTMATWIKPWEPLSVTLSEARSSSYNSDSQKAGPAVTHIRADYTLGHGFSVN  
ASVNNLFDITQAYSEGFEVGRNFWAGVEYTF"  
/complement (2426. .3280)  
/gene="ECs1694"  
/complement (2426. .3280)  
/gene="ECs1694"  
/note="molybdenum transport protein, similar to molybdenum  
transport proteins e.g. gi|3661478|gb|AAC61710.1| percent  
identity 91 in 284 aa"  
/codon\_start=1  
/evidences=not experimental  
/transl\_table=1  
/product="molybdenum transport protein"  
/protein\_id="BAB35117.1"  
/db\_xref="GI:13361159"  
/translation="MFLSQADIDALLLEDIQGGDLTTRALNIGHQHYEFFLRQGG  
CVSGVACKMLTTLGITDIDAVSDGQANAGRLIRAQNAALHQCWKAIONVLEW  
SCGVSDYLDQMLLRLRYPDGNIACRKAIPGTRLLASQAILAAGGLIHRAGCAETI  
LIFANHFHGDNDWGSALNOLREHAPEKKIVVEADTPKEAIALAQAQPNVLODKF  
SPQQAETIAQIAPSLAHCITLALTGGINLTYLTKNYLDGIRLFTSAPYAAPADIKV  
SLQPAASI"  
/complement (3219. .3500)  
/gene="ECs1695"  
/complement (3219. .3500)  
/gene="ECs1695"  
/note="unknown"  
/codon\_start=1  
/evidences=not experimental  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="BAB35118.1"  
/db\_xref="GI:13361160"  
/translation="MPTLILSVDDKIANKRITAPRNVLSRTSAGVLARLTMTSVSGYIAG  
INNKLVPSPAPATGSSGSIYARRRHCDPFPFSGTNRCVTSGRYPGR"  
/complement (3445. .4089)  
/gene="ECs1696"  
/complement (3445. .4089)  
/gene="ECs1696"  
/note="unknown, similar to hypothetical proteins e.g. Orf2  
[Escherichia coli CFT073] gi|3661479|gb|AAC61711.1|  
percent identity 98 in 214 aa"  
/codon\_start=1  
/evidences=not experimental  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="BAB35119.1"  
/db\_xref="GI:13361161"  
/translation="MLIDDDIDFADLYLQQLKLAHRTKTPDHWDQRAEKMAENCASPT  
DSYUQLIATIDLAQOITLFDMGCGPGTVSLADKLTYIGVDYSQGLMNVAAARRA  
ALKADNVHLTORAEEDWSLPRCDIIVASESTIVADMROAMSKLNNQARLVYTHL  
VSTSFVPAIQRAAGREVIEUPNIFALNVLYQMGIYAHVDFIRGNCQDQNSI"  
/complement (4099. .4857)  
/gene="ECs1697"  
/complement (4099. .4857)  
/gene="ECs1697"  
/note="probable ferric enterobactin transport, similar to  
ferric enterobactin transport ATP-binding protein  
[Escherichia coli CFT073] gi|3661480|gb|AAC61712.1|  
percent identity 79 in 148 aa"  
/codon\_start=1  
/evidences=not experimental  
/transl\_table=1  
/product="putative ferric enterobactin transport  
ATP-binding protein"  
/protein\_id="BAB35120.1"  
/db\_xref="GI:13361162"  
/translation="MTLIAVRHASLIGYSRHPVLRDVSFTLSQGTICCLLGANGCGKTT  
LMRSITGLVGLKGLDLPVOTLGHQRAQAIWPOAHGDGFAPFSDVLMGL  
APTIGAFVPGKQERLKATEOLEKLGILHLAARWNTLSGGEROLALIPALVQOPL  
LLLEDPASSLDFGHQILDLTAQLKNNGTMLMSTHPLJHANAIADSIIVQEPDGRV

gene  
CDS  
TQGLTQBLTNTKLAALRYVSADQIHHLHSAISH"  
/complement (4854. .5834)  
/gene="ECs1698"  
/complement (4854. .5834)  
/gene="ECs1698"  
/note="probable ABC transporter protein (permease), similar  
to ABC transporter permeases e.g. [Haemophilus influenzae]  
gi|2501391|sp|Q57130|YE71\_HAEIN percent identity 40 in 323  
aa"  
/codon\_start=1  
/evidences=not experimental  
/transl\_table=1  
/product="putative ABC transporter permease protein"  
Alignment Scores:  
Pred. No.: 5.96e-20 Length: 291136  
Score: 413.00 Matches: 88  
Percent Similarity: 62.50% Conservative: 27  
Best Local Similarity: 47.83% Mismatches: 53  
Query Match: 43.47% Indels: 16  
DB: 1 Gaps: 5  
US-10-009-916A-1 (1-180) x AP002556 (1-291136)  
Qy 1 MetLysLeuValLeuPheValThrSerIleValThrLeuSerLeuThrSerIle 20  
Db 276621 ATGAATGTAAATC-----ATTGTCGCAATGCGCATGCTGCACAGCA--- 276580  
Qy 21 ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle 36  
Db 276579 -----GCATCATCGGATACGACAGCAAGAGTCCCAATGAACCTTGTCT 276532  
Qy 37 AspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly 56  
Db 276531 AGTCTCAGCGAAAGAGTCAAGATTGGAATAATACCATTCAGGAGACCCCTACGGT 276472  
Qy 57 LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76  
Db 276471 CTGCTGTTCACACAGCGCCCTCACTCTGTCTGAAGGCATTCATGTTTTCATGTGAC 276412  
Qy 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96  
Db 276411 GAAAGAAATTCGCCCGCGCATCTGAAGACGGAACACCGCTCGCAGCAATTATCGGCT 276352  
Qy 97 HisGlyHisThrAspProAspLysThrGlyLysHisGlyGlyProLeuGly---AsnGly 115  
Db 276351 GCGGTCACITTTGACCGGAAACACCGGCAACATCTTGGCCCTGTCTCCGATGGA 276292  
Qy 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135  
Db 276291 CACCTGGCGGACCTCCCTCGCTGTCTGTCAGCATGACGGAAGGCAATACCCGCTC 276232  
Qy 136 LeuAlaProArgLeu---ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154  
Db 276231 CTGCCCCGAGACTGAACCTCAATAAAGAGATTAAAGGCGGTCTCTCATGCTTCAGCT 276172  
Qy 155 GlyGlyAspAsnThrSerAspLysProLeuProLeuGlyGlyGlyValAlaArgIleAla 174  
Db 276171 GCGGCTGATACCATCATGACCATCCGAGCCCTCGGCGTGTGTGTGCGAGAATGCGC 276112  
Qy 175 CysGlyValIle 178  
Db 276111 TCGCGCATCAATT 276100  
RESULT 35  
LOCUS AP005345/c 253750 bp DNA linear BCT 06-JUL-2004  
DEFINITION Vibrio vulnificus YJ016 DNA, chromosome II, complete genome,  
section 2/8.  
ACCESSION AP005345 BA000038  
VERSION AP005345.1 GI:37200415  
KEYWORDS Vibrio vulnificus YJ016

ORGANISM Vibrio vulnificus YJ016  
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
Vibrionaceae; Vibrio.

REFERENCE 1  
AUTHORS Chen, C.Y., Wu, K.M., Chang, Y.C., Chang, C.H., Tsai, H.C., Liao, T.L.,  
Liu, Y.M., Chen, H.J., Shen, A.B., Li, J.C., Su, T.L., Shao, C.P.,  
Lee, C.T., Hor, B.I., and Tsai, S.F.  
TITLE Comparative genome analysis of Vibrio vulnificus, a marine pathogen  
JOURNAL Genome Res. 13 (12), 2577-2587 (2003)  
PUBMED 14656965

REFERENCE 2 (bases 1 to 253750)  
AUTHORS Chen, C.Y., Wu, K.M. and Tsai, S.F.  
TITLE Direct Submission  
JOURNAL Submitted (29-MAY-2002) Shih Feng Tsai, National Health Research  
Institutes, Division of Molecular and Genomic Medicine; 128,  
Yen-Chiu-Yuan Road, Sec 2, Taipei 115, Taiwan  
(E-mail: petsai@nhi.org.tw, Tel: 886-2-8146-1041,  
Fax: 886-2-2789-0484)  
COMMENT This sequence was determined by the Sequencing Core of the National  
Yang-Ming University Genome Research Center (YMGRC;  
http://genome.ym.edu.tw).

FEATURES  
source  
1..253750  
/organism="Vibrio vulnificus YJ016"  
/mol\_type="genomic DNA"  
/strain="YJ016"  
/db\_xref="taxon:196600"  
/chromosomes="II"  
/gene="VVA0216"  
complement (125..1231)  
/gene="VVA0216"  
complement (125..1231)  
/note="identified by GeneMark and Glimmer2"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC96242.1"  
/db\_xref="GI:37200416"  
/translation="MYPVAKPASPMPIKMSASRSIKPSTSSWRMPLRCKSRPLP  
PRVRPOPGLVSSLLQKRNMTTATSSLLKSGKPKPLTETLLLLVAVFGTSGVL  
TKSALVYTSVLLFISFISITELCMLPVIRPQGLNKDWKIAIPTGFIILAIPECE  
VFGVSTASNAFLISLIVLPAPAEVLNKRISLILALVCGVIGVLLITSEQ  
IELSNTGDIFFILTAALLKALMVMTKRFTEGATITSTLSQSFVASCATVGVIV  
YVPAEIVLPSLEFVLVAVLVFCTLFVQYVAVRTSPTRVSLLMGSEPLFGA  
IFAMVWLQSLALQLLGGALILFSVIVSTRDN"  
complement (1032..1784)  
/gene="VVA0217"  
complement (1032..1784)  
/note="identified by GeneMark and Glimmer2"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC96243.1"  
/db\_xref="GI:37200417"  
/translation="MTMTSTHTKTKQVLPISQQLKKYKVFINGRQMTLLSTHVG  
EIVRFKNQWDRYMDGGTYRYYRQCTKLARSKEIVMMPEHPVQPSYINTL  
NGVERHFEPLDRFVTSVPLEKLMMSEIVDDVEGPQNRIHLHPYRITATAEA  
GPTPEGLHRDGVYIASMINKINVSGETLTDANKLELRIITLDKTFDVMADDT  
ATMWEVSTTPASRFSASRDVLIIFTKNEK"  
complement (1796..3010)  
/gene="VVA0218"  
complement (1796..3010)  
/gene="VVA0218"  
/codon\_start=1  
/transl\_table=11  
/product="argininosuccinate synthase"  
/protein\_id="BAC96244.1"  
/db\_xref="GI:37200418"  
/translation="MKKIRSIDVQNVASQAKHVLTLFSGGLDSSVLLILKNSRAKV  
TAVAILDDIDIESKLLITDHYGDLKILDAKQEVHSLFSAIQALYIGDPVVS  
SSLSRPIIVKAVELAQIGCDIAIHTANSQNSLRNLGAIERSGYQFGYSPSEYS  
AISREKAQTLFHSGLVGFKSRVNSGDSNLWCREFESGLDDPENFTLADSLFTWSRW

QPEHLEGDSIKIGFKQGYVPTINDKPIITLLELIAFINNHVGAYEIGHVYVGFHDLDQD  
EKVLEVRAPAAAMLKAYKLETTATLPTDVLKAKAHNETWTQEAVERGFGSLQNA  
SVAFIAHTAEHVSGSVLFLKSRGNALPTPSIVADNARYLDRDNWEINVAHARSWRVQ  
LEPTSDKHQVA"  
3115..4029  
/gene="VVA0219"  
3115..4029  
/gene="VVA0219"  
/codon\_start=1  
/transl\_table=11  
/product="transcriptional regulator"  
/protein\_id="BAC96245.1"  
/db\_xref="GI:37200419"  
/translation="MDTNRLIPLLSMAIFNVVSGSFKTAQKLGVSPSSVSRISIT  
RIENVLEKLLERTTROMLSATGORVYSLCSDMLNAAMVSAATDKSEASGSLRI  
APKALSROVLMPLVDFAEHPKVALQKADHYIDPIGDEVVDVLIQITDKPTEGLV  
ARSLGRVLCSADPYLQQQTPHPDDLVHNCCLGLENPRDNVMDFTIGSRKLSV  
NVKGSFVNHSEIRREAVLIGLISVFPPEFAIRPYIESGALTDELLSDMHLGGNYQGOI  
LAQYAKSKYVPOQIKTFIEYLQRMVHA"  
complement (4143..5447)  
/gene="VVA0220"  
complement (4143..5447)  
/gene="VVA0220"  
/codon\_start=1  
/transl\_table=11  
/product="proton/glutamate symporter"  
/protein\_id="BAC96246.1"  
/db\_xref="GI:37200420"  
/translation="MNTKPKMSLTGRVILGVAGILTGFAIRALPADNGFVDRIYVNG  
LFEVGKIFIASLKMVLVPLIFVSLVCGTSSLKDISTLGRMGKTLAFYITTTAVAIT  
LALTMTGTPQAGADLTAAASFSAEASLGQVIIDMFPTNPISAMAEKTLQTVIF  
ALLFGTASAGKPGERIAAIFSDLINEVIMKLVALIMNLAPYGVFFLMKLTGLGLG  
AIINLGEFLVLTALLHGLVTSYSLMKGFTGLNPLTFLKMWEDAIMFAFSPASSNA  
TIPATMETVKNLGVNDKIASFTVPLGATVNMDDGTALMOGVATFAAQAFNIDLSMGD  
YVMVILTATLASIGTAGVPGVGLMAMVLNQGLPLEGIALIMGVDRLLDMIRTAIVN  
ITGDACSVIVAKSEGLDLDRFPKAGEKEEVEHLKRADA"  
complement (5748..6848)  
/gene="VVA0221"  
complement (5748..6848)  
/gene="VVA0221"  
/note="identified by GeneMark and Glimmer2"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC96247.1"  
/db\_xref="GI:37200421"  
/translation="MMKNKRTLLILLASVSLSIAFAEVPAPVGRPCCAFGYDLKQAL  
GGIPVFFSVGNVDADKIGQHYNDGSSQVSGNLGLGDESNGLIFTQQAQFIDTAH  
VRDTADYTFYLYQENLAHLGQYQVELKOELRVQTHWQFQFQAESISQAEKVQSAE  
AAAFATQLAQWHEIAQWFGMSVGFDELASAFSPEDLYSNMLGAKLAKQVILTNPT  
LDKKTFSKNDRALQERLSKAKQSKITQOKIEALDGOWDWSARLPPKMLLKRDY  
HLGYTLQPNYPGSEHTQSLNEHFSQSPISDMVSLKLIASDEKHPHALPDLTKKQAV  
WTHQNPASLADPAPYQDARHHPKSPSPQVEE"  
complement (6855..7970)  
/gene="VVA0222"  
complement (6855..7970)  
/gene="VVA0222"  
/note="identified by GeneMark and Glimmer2"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC96248.1"  
/db\_xref="GI:37200422"  
/translation="MSKVGCCALLALTVSAPAKQSSWVDNLLKLGSSDVTDTSKL  
IDWGLPGPFRPEQGLGIGIAAGVLYPTGQSDADPYSTLTLLTASTSGSYGLGVE  
NRTYLCGRIRILLIDGWSHTPGYWGQESAQTDSSNKTQYNAQRYQNPXIAYEIA  
PSVYLKAGWHQTYQGVTVQDLNTPVAKLSNRSRSGMLIGMEYDFRDPENPAMGSY  
NLEWSDFRHGSQDNTYQDLNTPVAKLSNRSRSGMLIGMEYDFRDPENPAMGSY  
DDORVYRQGYQYRDNLSTQVEYHSHFNOPHGMVAVGACTIAPHYNELLEQSWLP  
TVGIGYRFAFRKARINVRVDMGFGKSSGSGFYFHVNEAF"  
complement (7974..8822)  
/gene="VVA0223"  
complement (7974..8822)

```
/gene="VVA0223"
/notes="identified by GeneMark and Glimmer2"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC96249.1"
/db_xref="GI:37200423"
/translation="MKKAVLLSLFAACSTPTVADSFSLPWKEDAEALGYTLPKPIG
PNSVMMEQGINVDSIDLOGLDFNYLDIGDAPRGQYDEVLTLRADWVLPFLNLV
PVLKDGSTVDLTGAFNPNRPFIHKLQDRLLDGYTTGFGVALVGGYENWEA
LVDSFGTSRUTVDIGTDAIVSVFRVGTDFNRHGTPLRIWAGAMYQDVEITLGSLS
DLGFPSSUTSLPSDARPEVQOHLQTPWNPVIGVQYQINESWLLYGFGRDSMFF
SVDRRF"
9024..9629
/gene="VVA0224"
9024..9629
/gene="VVA0224"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAC96250.1"
/db_xref="GI:37200424"
/translation="MAMKVTEQKRLALIEAAKDFIEPFGFNAANMDRVCEKASTSKRT
LYRHFSEKELFTRAINAWVEKQKHETFLYQAHVALDSQKAYLQKRVTLGYSEVGL
PVRMIVGEFIREPELAQRALMGDKDVKLQWTEINAIISGRILSGDSQOIMATLLN
LFHQGFVLPQIVANLSPNEQEQQAHHINEIVRIFSRAYVAD"
complement(9732..10202)
/gene="VVA0225"
complement(9732..10202)
/gene="VVA0225"
/notes="identified by Glimmer2"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC96251.1"
/db_xref="GI:37200425"
/translation="WDSFWSQLLYCLPNNLAHNNHVFGRGDFSRIRLALLGRVSS
NGYNTMTMTITSLALALAEKANTFRDQIFYNDGVHSHQVLGLNIKYEDFEIG
GYILRQETRYTLVHAADCIIEFWDFEQVRPESHKIVLPVHNSVGLPSLAQHNF"
10436..11338
/gene="VVA0226"
10436..11338
/gene="VVA0226"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAC96252.1"

Alignment Scores:
Pred. No.: 7,386-20 Length: 253750
Score: 411.00 Matches: 82
Percent Similarity: 63.53% Conservative: 26
Best Local Similarity: 48.24% Mismatches: 56
Query Match: 43.26% Indels: 6
DB: 1 Gaps: 2

US-10-009-916A-1 (1-180) x AP005345 (1-253750)

Qy 10 SerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerVal 29
Db 83111 ACCITACTGCGCCATTGTGTTATTCGACCTCTTCATTTGGCCAGAGTTTGAGCGTT 83052

Qy 30 ThrSerGluValHisMetIleAspAspAsnGlyIleLeysGlnSerIleGlyThrValThr 49
Db 83051 GAC-----ATGAAGACCTATCCAGCAATCAACACTCGGACCGTGACC 83007

Qy 50 PheThrAspThrAspLysGlyLeuGlnIleLeysThrAspLeuLeysGlyLeuProIadly 69
Db 83006 ATCAGTTCAAGTGACTACGGAACGGTGTGTTACTCTCTGATCTTAAGGGCTACCTAGTGGC 82947

Qy 70 GluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHis 89
Db 82946 CTGCACGGTTTCATCTCCACGCCCAATGAAGCTGTGAGCAGCAACAAAGACGGCAAA 82887

/gene="VVA0223"
/notes="identified by GeneMark and Glimmer2"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC96249.1"
/db_xref="GI:37200423"
/translation="MKKAVLLSLFAACSTPTVADSFSLPWKEDAEALGYTLPKPIG
PNSVMMEQGINVDSIDLOGLDFNYLDIGDAPRGQYDEVLTLRADWVLPFLNLV
PVLKDGSTVDLTGAFNPNRPFIHKLQDRLLDGYTTGFGVALVGGYENWEA
LVDSFGTSRUTVDIGTDAIVSVFRVGTDFNRHGTPLRIWAGAMYQDVEITLGSLS
DLGFPSSUTSLPSDARPEVQOHLQTPWNPVIGVQYQINESWLLYGFGRDSMFF
SVDRRF"
9024..9629
/gene="VVA0224"
9024..9629
/gene="VVA0224"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAC96250.1"
/db_xref="GI:37200424"
/translation="MAMKVTEQKRLALIEAAKDFIEPFGFNAANMDRVCEKASTSKRT
LYRHFSEKELFTRAINAWVEKQKHETFLYQAHVALDSQKAYLQKRVTLGYSEVGL
PVRMIVGEFIREPELAQRALMGDKDVKLQWTEINAIISGRILSGDSQOIMATLLN
LFHQGFVLPQIVANLSPNEQEQQAHHINEIVRIFSRAYVAD"
complement(9732..10202)
/gene="VVA0225"
complement(9732..10202)
/gene="VVA0225"
/notes="identified by Glimmer2"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC96251.1"
/db_xref="GI:37200425"
/translation="WDSFWSQLLYCLPNNLAHNNHVFGRGDFSRIRLALLGRVSS
NGYNTMTMTITSLALALAEKANTFRDQIFYNDGVHSHQVLGLNIKYEDFEIG
GYILRQETRYTLVHAADCIIEFWDFEQVRPESHKIVLPVHNSVGLPSLAQHNF"
10436..11338
/gene="VVA0226"
10436..11338
/gene="VVA0226"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAC96252.1"

Alignment Scores:
Pred. No.: 7,386-20 Length: 253750
Score: 411.00 Matches: 82
Percent Similarity: 63.53% Conservative: 26
Best Local Similarity: 48.24% Mismatches: 56
Query Match: 43.26% Indels: 6
DB: 1 Gaps: 2

US-10-009-916A-1 (1-180) x AP005345 (1-253750)

Qy 10 SerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerVal 29
Db 83111 ACCITACTGCGCCATTGTGTTATTCGACCTCTTCATTTGGCCAGAGTTTGAGCGTT 83052

Qy 30 ThrSerGluValHisMetIleAspAspAsnGlyIleLeysGlnSerIleGlyThrValThr 49
Db 83051 GAC-----ATGAAGACCTATCCAGCAATCAACACTCGGACCGTGACC 83007

Qy 50 PheThrAspThrAspLysGlyLeuGlnIleLeysThrAspLeuLeysGlyLeuProIadly 69
Db 83006 ATCAGTTCAAGTGACTACGGAACGGTGTGTTACTCTCTGATCTTAAGGGCTACCTAGTGGC 82947

Qy 70 GluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHis 89
Db 82946 CTGCACGGTTTCATCTCCACGCCCAATGAAGCTGTGAGCAGCAACAAAGACGGCAAA 82887
```



DTRELNDPQVPPRIEGLKLEPVGMQGGQORVIAGLEVTIVVEQSFRLTABQPGTFTLL  
PQITGSIYVGDSTIKTQKMPITVKIAPSEFKSGVLPGLSALQMTQSQW  
DDGNTANTVANNVKGSSITQIKQAGTQAEYLPRTIMDYPNLSRVVPEQPDF  
TARDGTVIMTKQVLIITEAGEFTLPQYTLNWDKSKDEAKQANLSKLNVQSDAG

## Alignment Scores:

Pred. No.: 1.15e-19 Length: 335862  
Score: 410.00 Matches: 81  
Percent Similarity: 65.00% Conservative: 23  
Best Local Similarity: 50.62% Mismatches: 50  
Query Match: 43.16% Indels: 6  
DB: 1 Gaps: 2

US-10-009-916A-1 (1-180) x AP005089 (1-335862)

Qy 20 IleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsn 39  
Db 79365 ATGACCAACACCGCTGGCCCAATCAATCAC-----ATGCCGATTTG 79409  
Qy 40 GlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIle 59  
Db 79410 GGTACTAACCAACCGCTGGCTACGCTGACTTATCAGAGAGTGACTACGGAATGTATTC 79469  
Qy 60 LysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlyGly 79  
Db 79470 ACGCGAATCTGTCCAGAAATCTCTCGCGGACTACATGCTTTTCCGTCACCGCTAACCCCA 79529  
Qy 80 SerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHis 99  
Db 79530 TCCGTGACAGTGCAGAAAGACGGAACCGTAAACCTAGTAGTGGCGGTGGCGGCCAT 79589  
Qy 100 TyrAspProAspLysThrGlyLysHisGlyGlyProLeuGlyAsnGly---HisLysGly 118  
Db 79590 TACGACCAGACATACCTGGCAACATGTTTCCATGGACCGATGGCAACCATCTTGGC 79649  
Qy 119 AspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaPro 138  
Db 79650 GATCTGCAGCGTTGTACGCGGACATGAGGAAACCGCAACCGCTTTTACGCTCG 79709  
Qy 139 ArgLeuThrValLysGluIleLysGlyArgThrValMetIleHisAlaGlyLysAspAsn 158  
Db 79710 CGCTCAAAATGTCGGATTAAAGGACGACCTTGATGATCATGCTGGCGGTGATAT 79769  
Qy 159 TyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
Db 79770 CACTCGGATCAACGACCAATTAGTGGCGGCTGCACGCATTTGTGTGGCGTGATC 79829  
RESULT 37  
AE016813/c 336182 bp DNA linear BCT 29-SEP-2003  
LOCUS Vibrio vulnificus CMCP6 chromosome II section 6 of 6 of the  
DEFINITION complete sequence.  
ACCESSION AE016813 AE016796  
VERSION AE016813.1 GI:27359304  
KEYWORDS Vibrio vulnificus CMCP6  
SOURCE Vibrio vulnificus CMCP6  
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
Vibrionaceae; Vibrio.  
REFERENCE 1 (bases 1 to 336182)  
AUTHORS Kim,Y.R., Lee,S.E., Kim,C.M., Kim,S.Y., Shin,E.K., Shin,D.H.,  
Chung,S.S., Choy,H.B., Proguiske-Fox,A., Hillman,J.D., Handfield,M.  
and Rhee,J.H.  
TITLE Characterization and Pathogenic Significance of Vibrio vulnificus  
Antigens Preferentially Expressed in Septicemic Patients  
JOURNAL Infect. Immun. 71 (10), 5461-5471 (2003)  
PUBMED 14500463  
REFERENCE 2 (bases 1 to 336182)  
AUTHORS Jeong,H., Moon,Y.H. and Kim,J.J.  
TITLE Direct Submission  
JOURNAL Submitted (13-DEC-2002) Genotech Corp., 461-6, Jeonmin-dong,  
Yuseong-gu, Daejeon 305-811, South Korea  
REFERENCE 3 (bases 1 to 336182)

AUTHORS  
TITLE  
JOURNAL

Rhee,J.H., Kim,S.Y., Chung,S.S., Lee,S.E. and Choy,H.E.  
Direct Submission  
Submitted (13-DEC-2002) Department of Microbiology, Genome Research  
Center for Enteropathogenic Bacteria, Chonnam National University  
Medical School, Hak-1 Dong, Dong-Gu, Kwang-Ju 501-746, South Korea

FEATURES  
source

1. 336182  
/organism="Vibrio vulnificus CMCP6"  
/mol\_type="genomic DNA"  
/strain="CMCP6"  
/db\_xref="taxon:216895"  
/chromosome="II"  
288..1034  
/locus\_tag="VV21363"  
288..1034  
/locus\_tag="VV21363"  
/note="COG0583"  
/codon\_start=1  
/transl\_table=11  
/product="Transcriptional regulator"  
/protein\_id="AAO08249.1"  
/db\_xref="GI:27359305"  
/translation="MAERICLTQPGVSHVKKLEYLLGSKVIKRGKKIETLHGQLL  
YDANDLRKMEHLNKLNGYISVPISFSSDCLSTIEKLEELKIGYESSVIVH  
GNNVQTADVQLTPVFGDKTISLSEFHNFIQLGRDSDSDPDIETVMYSEPLKRVV  
REITHNGYKIAETCRWIPMGSRIMQYTHMNGRSLMVCPDWQNSQFAVSTPLDIOV  
RFLMKVTHDIGHLDTAKTQLNVLTIESLFTG"  
complement (1101..1934)  
/locus\_tag="VV21364"  
complement (1101..1934)  
/locus\_tag="VV21364"  
/note="COG0005"  
/codon\_start=1  
/transl\_table=11  
/product="Purine nucleoside phosphorylase"  
/protein\_id="AAO08250.1"  
/db\_xref="GI:27359306"  
/translation="MTEDPVPNQALETIYSYKPNYQPKVAFILGSLGLVIADELDDKIV  
IQEFLGFFSVTVEGSHGLVLEKLFIDVICMKGRHYEHQTMKVTNPVTFKR  
LGCFELVTTNAGLSRDERIDVGLVFNVDHINTPESPMTPNDNDAYGPRFFSLANA  
VLSRSEALTVAEAGIHNLQGVFVSYTGNFTAAEIRMMQIIGDVGWGVVPEV  
ISAAHCSLPVALCAITNMAEGLSDVTLSHAQTLACAKLAEDAFIKLIKAFIKSHEAT  
S"  
2072..2944  
/locus\_tag="VV21365"  
2072..2944  
/locus\_tag="VV21365"  
/note="COG0583"  
/codon\_start=1  
/transl\_table=11  
/product="Transcriptional regulator"  
/protein\_id="AAO08251.1"  
/db\_xref="GI:27359307"  
/translation="MLRYDFKVAQLGQFSRAAEQLNITKPSLQIKELPEVLGTALF  
ERNTRVQLTAGEQLQVECVTLPAQFERSIHRVKQCARERQQLDGLMSISIFWAGF  
GPAFHQLQMLPHATLNLLELSPKQVRQILMHQIDGLVRFADRVAPLLSLCYO  
ECMVVLPSEHPLAQGLTSLKADFVMLKQENSSSTLRIAEYCAKAGYRNIVQ  
EVEPNTLLAVISARQAVSIVPSQYANLSWPQVFIEPLKQIPAGVVALYHPEAQSN  
KRMIDELKQALEK3"  
complement (3014..5149)  
/locus\_tag="VV21367"  
complement (3014..5149)  
/locus\_tag="VV21367"  
/note="exoprotease Vcc"  
/codon\_start=1  
/transl\_table=11  
/product="Conserved hypothetical protein"  
/protein\_id="AAO08252.1"  
/db\_xref="GI:27359308"  
/translation="MTLQQNKRSLALAIIFGLSGCGDGGSSSSSTVQPKDPTPT  
TNPDRTKSGQCTATOLIADTSFASQIANADYDCMNSWFTSQIEADALFSEQNLA  
LVNDALLETVNOYGEGETAKHIVYLAETKAAVKNRHDYVARKIAPSTLSQNIAT  
TSATFLRNTHALENGDEQMKALGSMILLIVDSVRQLHIAASDVFAALLNQFEQLTENY









```

gene      complement(6167. .6532)
CDS       /gene="y0817"
          complement(6167. .6532)
          /gene="y0817"
          /function="putative"
          /notes="residues 2 to 121 of 121 are 82.49 pct identical to
          residues 1 to 120 of 121 from E. coli K12 : B2765"
          /codon_start=1
          /transl_table=11
          /product="putative 6-pyruvoyl tetrahydrobiopterin
          synthase"
          /protein_id="AA084404.1"
          /db_xref="GI:21957547"
          /translation="MIMSTTLFLKDFQFEAAHRLPHVPEGHKGCRLHGHFMYRLEVTG
          EYDAHSQWMDFAALKAQVPIWRLDHHVLDIPLGNPTSEVLARIWVQLKPKQLP
          ELASVMIKETCSAGCVYRG"
gene      6877. .8757
CDS       /gene="cysJ"
          /note="synonym: y0818"
          6877. .8757
          /gene="cysJ"
          /function="enzyme; central intermediary metabolism: Sulfur
          metabolism"
          /notes="residues 15 to 626 of 626 are 73.40 pct identical
          to residues 1 to 599 of 599 from E. coli K12 : B2764"
          /codon_start=1
          /transl_table=11
          /product="sulfite reductase (NADPH), flavoprotein beta
          subunit"
          /protein_id="AA084405.1"
          /db_xref="GI:21957548"
          /translation="MCQPLKHKHGIRMTTQAPTSLPLSPEQLARLQAAGFERSP
          TOMALSGYFMWQNGQPGAVAPPAVPVTVLISASQTGNARRLAEQLRDLILA
          AOLSVNLVAGDYKFOIAQERLLVVVASTQGEGRPAERAVLHKFLFSKAPKLSET
          AFAVPLGTSVEHFCQAGKDFSKLAELGAQRLDRVDADVEYQVQAQWQOVVAT
          LQAVPAQTATQFTAPTQPTPAAALITSGTITTVSPYKAPLTAQLSVQQRVTG
          RNSEKDRHIEDLGSGLRYQFDGALVWFNDPALVELLALLWLKGDPEVSDIGQ
          NNPLAQLSHLELQNTLLIYDKVYALSRDETLIALADKPAQLIYAKNTPFVDMVR
          QAPSLNAQLVLLRPLRLYSIASQAQATENEVHTVGVVRYDIDGRASGGASG
          VIADRLVDPGDIRVLEHNDRLPANPETVPMIGPGIAPFRAPMQGRVDGASG
          KWLPLFGNHFTEFLYQVWQRYKVEGVLTIDLAWSRDQAKHIVYQKLREQGREL
          MWMIQGAHLYVCGDANRWAKQVEVLDDVALHGAADAEQADEYLSLRLQRRYQRD
          VT"
gene      8757. .10487
CDS       /gene="cysI"
          /note="synonym: y0819"
          8757. .10487
          /gene="cysI"
          /function="enzyme; central intermediary metabolism: Sulfur
          metabolism"
          /notes="residues 1 to 571 of 576 are 80.38 pct identical to
          residues 1 to 570 of 570 from E. coli K12 : B2763;
          residues 1 to 571 of 576 are 80.91 pct identical to
          residues 1 to 570 of 570 from GenPept : >gb|AA021827.1|
          (AE008834) sulfite reductase, alpha subunit, NADPH
          dependent [Salmonella typhimurium LT2]"
          /codon_start=1
          /transl_table=11
          /product="sulfite reductase, alpha subunit"
          /protein_id="AA084406.1"

```

## Alignment Scores:

```

Pred. No.: 1.15e-20 Length: 11361
Score: 404.50 Matches: 92
Percent Similarity: 52.56% Conservative: 21
Best Local Similarity: 42.79% Mismatches: 63
Query Match: 42.58% Indels: 39
DB: 1 Gaps: 4

```

US-10-009-916a-1 (1-180) x AE013684 (1-11361)

```

QY      3 IleLysLeuPheValThrSerIle-----Val 12
      ::||| ||| ::||| ::

```

```

Db      4508 GTCAAAAGTTTCTATTTGAAGGTTTAAACGACACTTAATCAACATATAGGGATATAAC 4567
QY      13 ThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerValThrSerGlu 32
      ::||| ||| ::||| ||| ::||| ||| ::||| ||| ::||| |||
Db      4568 ATATGAAGTTAAGTACATTATGCTACCTGTTATTTCTTACTCCAGCGCAACACTGGCC 4627
QY      33 ValHisMetIleAspAspAsn----- 39
Db      4628 GCTAACATCGTGGCATGAATGATAGGCGCAGCATGAATGATAAGGCCAGCATGAATGGT 4687
QY      40 -----GlyIleLysGlnSerIle 45
Db      4688 AAGGCTAGCATGACTGTGAAATCAACGAATCATTCGCCACAAGGTAAATGGAAAGCGCTT 4747
QY      46 GlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGly 65
Db      4748 GGCACCGTACCGTGACTGAAACCGCTTATGGCTTACTGTTTACGCCACATCTCATCTGGG 4807
QY      66 LeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGlu 85
Db      4808 CTGGCTCCGGGAATTCACGGTTTCCATCTGCATGATAAAACCCAGTTGTGCTCGGGGATG 4867
QY      86 HisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisIleThrAspProAspLysThr 105
Db      4868 AAAGATGGCAAGCAGTGCAGCATTTGGCAGCGCGGGGGGCATCTTGCACCAATAAGACC 4927
QY      106 GlyLysHisGluGlyProLeuGlyAsn---GlyHisLysGlyAspLeuProArgLeuVal 124
Db      4928 GGGGTACACTTGGTCTTACACGATAAAGGCGCATCTGGGGATCTGCGGGATCTGGGATTTGGT 4987
QY      125 ValLysAlaAspGlyLeuAlaLysGluThrLeuLeuAlaProArgLeu---ThrValLys 143
Db      4988 GTTAATGCAGATGGCACCACCCTATCCCGTATTGGTCCGCGCCTGAAATCGTTGTGCA 5047
QY      144 GluLeLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLysPro 163
Db      5048 GAGGTGAAACAGCATGCGTTAATGATGCATGCTGGCGGTGATAATTACTTGTATCATCCA 5107
QY      164 LeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db      5108 ATGCCTTTAGCGGTGGTGGCGCAGCATGGCATGGCATGGAGTGCATT 5152

RESULT 39
AJ414156/c
LOCUS   AJ414156.1 220050 bp DNA linear BCT 21-NOV-2002
DEFINITION Versinia pestis strain CO92 complete genome; segment 16/20.
ACCESSION AJ414156 AU590842
VERSION   AJ414156.1 GI:15981150
KEYWORDS .
SOURCE    Versinia pestis CO92
ORGANISM  Versinia pestis CO92
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Versinia.
REFERENCE 1 (bases 1 to 220050)
          Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T.G.,
          Preteille, M.B., Sebaihia, M., James, K.D., Churcher, C., Mungall, K.L.,
          Baker, S., Basham, D., Bentley, S.D., Brooks, K., Cerdeno-Tarraga, A.M.,
          Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Dougan, G.,
          Fellwell, T., Hamlin, N., Holroyd, S., Jagels, K., Leather, S.,
          Karylshev, A.V., Moule, S., Oyston, P.C.F., Quail, M., Rutherford, K.,
          Simmonds, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G.
          Genome sequence of Versinia pestis, the causative agent of plague
          Nature 413 (6855), 523-527 (2001)
          21470413
          11586360
REFERENCE 2 (bases 1 to 220050)
          Parkhill, J.
          Direct Submission
          Submitted (04-OCT-2001) Submitted on behalf of the Versinia
          sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
          Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
          Notes:
          Details of Y. pestis sequencing at the Sanger Centre are available

```

on the World Wide Web.  
(URL, [http://www.sanger.ac.uk/Projects/Y\\_pestis/](http://www.sanger.ac.uk/Projects/Y_pestis/)).

## FEATURES

## source

Location/Qualifiers

1..220050

/organism="Yersinia pestis CO92"

/mol\_type="genomic DNA"

/strain="CO92"

/db\_xref="taxon:214092"

/note="biovar: Orientalis"

complement(191..715)

/gene="arol"

/note="synonym: YPO3215"

complement(191..715)

/gene="arol"

/EC number="2.7.1.71"

/note="Similar to Escherichia coli shikimate kinase II

AroL SW:AROL ECOLI (P08329) (174 aa) fasta scores: E():

1.5e-31, 55.4% id in 168 aa, and to Erwinia chrysanthemi

shikimate kinase AroL or AroM SW:AROL ERWCH (P10880) (173

aa) fasta scores: E(): 0, 63.7% id in 168 aa"

/codon\_start=1

/transl\_table=11

/product="shikimate kinase II"

/protein\_id="CAC92450.1"

/db\_xref="GI:15981151"

/db\_xref="GOA:Q8ZC15"

/db\_xref="Swiss-Prot:Q8ZC15"

/translation="MTQTFMVGARGAGKTTIGKALQALGYRFVDTDLFMQOTSQMT

VAEVESGWDGFLRESMALQAVTAPKVTATGGAVLSSENAFMEDHGRVLYLRA

SAAVLAKLADPEAQRPSLTGPIVEILDVLASREALYQDVAAHVLDGTQTPSLV

VEQIIQLMUTGEWK"

complement(209..709)

/gene="arol"

/note="Pfam match to entry PF01202 SKI, Shikimate kinase,

score 245.00, E-value 1.1e-69"

complement(470..544)

/gene="arol"

/note="PS01128 Shikimate kinase signature."

complement(1668..691)

/gene="arol"

/note="PS00017 ATP/GTP-binding site motif A (P-loop)."

complement(1962..1519)

/gene="arol"

/note="PS03216"

complement(1962..1519)

/gene="YPO3216"

/note="Similar to Rhizobium meliloti phospholipid

N-methyltransferase PmtA TR:AAG10237 (EMBL:AF201695) (200

aa) fasta scores: E(): 1.1e-07, 25.9% id in 185 aa, and to

Bradyrhizobium japonicum PmtA protein TR:Q9LCT2

(EMBL:Y09633) (199 aa) fasta scores: E(): 0.0057, 24.2% id

in 178 aa, and to Klebsiella pneumoniae hypothetical 22.7

kDa protein TR:O87755 (EMBL:AJ011907) (196 aa) fasta

scores: E(): 0, 52.9% id in 191 aa"

/codon\_start=1

/transl\_table=11

/product="putative methyltransferase"

/protein\_id="CAC92451.1"

/db\_xref="GI:15981152"

/db\_xref="GOA:Q8ZC14"

/db\_xref="TrEMBL:Q8ZC14"

/translation="MLNIVRRNFKSQPSYLORFIASPTVGSGLVPSSPWLQAMLNQI

DWTQNLTAELGAADGVLTTRILGHMSANLQAYETQPHFVHALHQLKDSRLQIADR

SAEQDQYDVVFCCLPISPTKTSIRILOQTQQLRARNGVLVLFQYSHLSBSLLS

RYFNWKKIRVVRNFPALVYICHN"

1787..2002

/gene="YPO3217"

1787..2002

/gene="YPO3217"

/note="no significant database hits"

/codon\_start=1

/transl\_table=11

/product="hypothetical protein"

/protein\_id="CAC92452.1"

/db\_xref="GI:15981153"

/db\_xref="TrEMBL:Q8ZC13"

/translation="MIGTDVGIVVANVIMARTNDRGRISCHFFIOSSPNAHPLNVD

IKNIASPHVLSIVACSIKKAAXKRLS"

complement(2249..2770)

/gene="YPO3218"

complement(2249..2770)

/gene="YPO3218"

/note="Similar to Escherichia coli hypothetical protein

YafE SW:YAFE ECOLI (P30866) (207 aa) fasta scores: E(): 0,

65.1% id in 169 aa, and to Pseudomonas aeruginosa YafE or

PA3119 TR:O87011 (EMBL:U93274) (187 aa) fasta scores: E():

0, 55.2% id in 172 aa"

/codon\_start=1

/transl\_table=11

/product="conserved hypothetical protein"

/protein\_id="CAC92453.1"

/db\_xref="GI:15981154"

/db\_xref="GOA:Q8ZC12"

/db\_xref="TrEMBL:Q8ZC12"

/translation="MQAAADKKLTNIEIQQLAELSPADQSFDTIVISYSAHHWDV

GKALREVKRVLPGGKVFMDVSPGHPVLDIYLTQVEVLDTGHVRNYSPEGLSLF

TEGLVINEVTSDSLYLEFSSWVARMTPAHFATAREFQKLSADSGVINHVAIQADGS

FTSDIMMIVARD"

complement(3016..3285)

/gene="YPO3219"

complement(3016..3285)

/gene="YPO3219"

/note="no significant database matches"

/codon\_start=1

/transl\_table=11

/product="putative membrane protein"

/protein\_id="CAC92454.1"

/db\_xref="GI:15981155"

/db\_xref="TrEMBL:Q8ZC11"

/translation="MEGSMGISEERSIRRLTNEKNVGHSAKWVAISAVYFILMF

YKHEIGVLTLAGGIFLVSTFTMKRKQKVKYKQQLQOIEEDKTV"

complement(3088..3138)

/gene="YPO3219"

/note="One of 2 probable transmembrane helices predicted

for YPO3219 by TMHMM2.0"

complement(3151..3216)

/gene="YPO3219"

/note="One of 2 probable transmembrane helices predicted

for YPO3219 by TMHMM2.0"

complement(3468..4226)

/gene="YPO3220"

complement(3468..4226)

/gene="YPO3220"

/note="Similar to Escherichia coli hypothetical protein

Ycgr SW:Ycgr ECOLI (P76010) (244 aa) fasta scores: E(): 0,

40.7% id in 241 aa"

/codon\_start=1

/transl\_table=11

/product="conserved hypothetical protein"

/protein\_id="CAC92455.1"

/db\_xref="GI:15981156"

/db\_xref="TrEMBL:Q8ZC10"

/translation="MGFVSETSKQFVKKNKLAICAILDLQKNDTAVMVTHARGQFI

SRILDIQETNQFIPIDFGSVENENVALGARQLIIVPTGAKIEFTCNKIKHVEYLS

LPATSSAIPQOLYIQRREYFRVPIQWPATYKSGKPDGQYKYLADIISGGNGLY

AMKSGEPFLQCSVLDAADVLCGFLFKDLQFIRALDKQVNVNKGPMITVQRLSFK

FLRLSPIQEGLQRAIFELEKQQTAKARKFOEGL"

complement(4554..5813)

/gene="proA"

/note="proA"

/note="synonym: YPO3221"

complement(4554..5813)

/gene="proA"

/EC number="1.2.1.41"

/note="Similar to Serratia marcescens gamma-glutamyl

phosphate reductase ProA SW:PROA SERMA (P17857) (417 aa)

fasta scores: E(): 0, 85.1% id in 417 aa, and to

Escherichia coli gamma-glutamyl phosphate reductase ProA



```

7,8-Dihydro-6-hydroxymethylpterin-pyrophosphokinase, HPPK
/interpro|IPR000550, (GO:0009396))"
/codon_start=1
/transl_table=11
/product="2-amino-4-hydroxy-6-
hydroxymethylidihydropteridine pyrophosphokinase"
/protein_id="AAS60560.1"
/db_xref="GI:45435000"
/translation="MIRVIALGSLNPLQOVSAAREALAHLPRLSRVLVAGSPLYRTK
PLGQDQDFENAVVALDTSLPPEQLDHTQALPENQGRVKEQSGWGPRTLDLDMLY
GEQVKTDRILIPHYGLKAREFMLYPLADIAFDLIPDGSLSSECLKRVXNGLVIM"
/complement(2809..4230)
/gene="pcnB"
/locus_tag="YP0286"
/complement(2809..4230)
/gene="pcnB"
/locus_tag="YP0286"
/note="highly similar to Yersinia pestis CO92 pcnB and
Yersinia pestis KIM pcnB. Contains Polynucleotide
adenyllyltransferase (Interpro|IPR002646, (GO:0006396))"
/codon_start=1
/transl_table=11
/product="poly(A) polymerase"
/protein_id="AAS60561.1"
/db_xref="GI:45435001"
/translation="MPGSDNVARKERRPARAHTGKHAVSSSEQRQWAIIPRDQHNI
SRDTSNDPLKVLRLNKSQYEAIVGGVDRDLILGRKPDFDITTSATPEQVKLFR
NCRVGRFRLAHVMGPEIIEVATFRGHEQOQAEQSDKNSQOQNGMLLRDNI FG
SIEDDQREDTINSLYGIDSFALRDYTGGLRDLEKGIILRIGDPETRYEDVRML
RAVRPAKLMSISPETAPRPLASLRPEIPPARLFESLKLQSGYGYKYLKCE
YQLFPLEPLARNTPEQDSHMERILVQKNTDRLHNDQVRNPAFLFAAMLWYPL
IEHAQKLTQESGLAYYDAFAMNDVLDDECSRSLAIPKRIITSLVDIWLQLRLSRQ
GKRAHLMHPKFAAYDILLIRAEVKNHQLRLAQWQGFQEAFTPQQKSMNLTLG
ADAPRRSRPRPRKVPVKGV"
/complement(4373..5527)
/gene="glnS1"
/locus_tag="YP0287"
/complement(4373..5527)
/gene="glnS1"
/locus_tag="YP0287"
/note="highly similar to Yersinia pestis CO92 YPO3398 and
Yersinia pestis KIM Y0789. Contains Glutamyl-tRNA
synthetase, class Ic (Interpro|IPR000924, (GO:0006424))"
/codon_start=1
/transl_table=11
/product="putative glutamyl-tRNA synthetase"
/protein_id="AAS60562.1"
/db_xref="GI:45435002"
/translation="MLEDTGFORLFSSQLFDPQPGSKPSGSQLFGSQPFPAALQFLSA
AQPRASQPSAQLPETQPVVQQAIVIQSANRATNPTEYIGRFAPSPGDLHFGSL
IAALGSYLQARAQGMVLRIEDIDPPREVFGAASRIILAEHLVGLHWGDFVYQSOR
HEAVRATNLWLEQOGLSYCTCTRSRIHQGLGYDGYCRDRHLPASGAAILRLOTQPV
YAFYDKLIGELHAPALAQEDFIIRRDGLFAYNLAVVDDAFQGVTEIVSGADLIEP
TVRQIALYQQQLCHPFGYHILPLALNNQGNKLSKQNHAPPPLNGDPRIILDALKFLR
QPLRYWQDLDLIRLFAVEHWLTVLSIPLOQAITPQKTRHSQSKYGL"
/5419..5448
/note="degenerate repeat region containing 2 copies of a
15bp unit"
/rpt_type=tandem
5579..5612
/note="degenerate repeat region containing 2 copies of a
17bp unit"
/rpt_type=tandem
/complement(5618..6166)
/gene="dksA1"
/locus_tag="YP0288"
/complement(5618..6166)
/gene="dksA1"
/locus_tag="YP0288"
/note="highly similar to Yersinia pestis CO92 dksA and
Yersinia pestis KIM dksA. Contains Zn-finger, prokaryotic
DksA/Trar C4 type (Interpro|IPR000962))"
/codon_start=1

```

```

/transl_table=11
/product="Dnak suppressor protein homologue"
/protein_id="AAS60563.1"
/db_xref="GI:45435003"
/translation="MPYILPVSIIYGLIPIPIGDFDSAYECRRSIMQSGOKRKTSSLS
ILAIAGVEPYQKEGEEYMNAAQSHFKLILEARNQRURDEVDRTVSHMQDEAANFPD
PVDRAAQEEFSELELRNDRERKLIKIKTEKTKVEDDDFGFCSCGVEIGIRLEAR
PTADLCIDCKTLAEIREKQAG"
/complement(6243..7118)
/gene="sfsA"
/locus_tag="YP0289"
/complement(6243..7118)
/gene="sfsA"
/locus_tag="YP0289"
/note="highly similar to Yersinia pestis CO92 sfsA and
Yersinia pestis KIM sfsA. Contains Sugar fermentation
stimulation protein (Interpro|IPR005224)"
/codon_start=1
/transl_table=11
/product="sugar fermentation stimulation protein"
/protein_id="AAS60564.1"
/db_xref="GI:45435004"
/translation="MKGQMPANANLPFSANLLSVNSDSRSISPNSPPTHAKSLQFN
APLQPTPQATLILIRKRFADIVTPAGEALTIHCANTGAMTCATPGDTIWTST
SDNPKRYQPSQWELTQTQDMICVNTWRANELVNLAIEKQIABLSGYNFVRKEVY
GEENSRIDLLQAEDRRDCYTEVKSVTLLQQCGYFPDAVTLRGQHLRELNQRYVNG
HRAVLPFAVLHTGIKQVAPARHIDRRVRAELLVQAQAGVEVICYGFQLSPDGIENLR
LPLLDLMLSENAAE"
/complement(7093..7668)
/gene="ligT"
/locus_tag="YP0290"
/complement(7093..7668)
/gene="ligT"
/locus_tag="YP0290"
/note="highly similar to Yersinia pestis CO92 ligT and
Yersinia pestis KIM Y0793. Contains 2'_5' RNA ligase
(Interpro|IPR004175, (GO:0016070))"
/codon_start=1
/transl_table=11
/product="putative 2'-5' RNA ligase"
/protein_id="AAS60565.1"
/db_xref="GI:45435005"
/translation="MTNTVDBELKVRRLFPALTLPDMQOEIVQWRAGHFSPGGRPV
AAANHLTLAFLEGVSATKSQLQOQAGRIQQAGFSVTLDIDIGHWPGSGVTLGCKNP
PRGLLQALRLSQAARSQCYQTPLPHPHVTLLRNATRPVALPAKSGNETFQADHFS
LYESVFRAGRTRYNIQSWPLAGSERKPDAC"
/7709..10270
/gene="hrpB"
/locus_tag="YP0291"
/7709..10270
/gene="hrpB"

```

Alignment Scores:

| Pred. No.:             | 2,83e-19 | Length:       | 290803 |
|------------------------|----------|---------------|--------|
| Score:                 | 404.00   | Matches:      | 83     |
| Percent Similarity:    | 66.01%   | Conservative: | 18     |
| Best Local Similarity: | 54.25%   | Mismatches:   | 50     |
| Query Match:           | 42.53%   | Indels:       | 2      |
| DB:                    | 1        | Gaps:         | 2      |

US-10-009-916A-1 (1-180) x AB017128 (1-290803)

| QY | 28    | SerValThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThr    | 47    |
|----|-------|--|-------|
| Db | 37285 | AGCATGACTGTGAAATCACAAGTATCGCCACAGGTAATGGGAAGCGCTTGGCACC      | 37344 |
| QY | 48    | ValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuPro | 67    |
| Db | 37345 | GTGACGGTGAAGTAAACCGCTTATGGCTTACTGTTTACGCCACATCTCCTGGCGTGGCT  | 37404 |
| QY | 68    | AlaGlyGluHisGlyPheHisIleHisGlyGlySerCysGlyProAlaGluHisAsp    | 87    |
| Db | 37405 | CCGGGAATTCACGGTTTCATCTGCATGAAACCAACCACTTGTGCTCCGGGGATGAAGAT  | 37464 |



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: November 1, 2004, 05:38:33 ; Search time 64 Seconds  
(without alignments)  
1999.094 Million cell updates/sec

Title: US-10-009-916A-1  
Perfect score: 950  
Sequence: 1 MKIKLFFVTSIVTISLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues  
Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US1000916/runat\_26102004\_100224\_2000/app\_query.fasta\_1.327  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1000916 -CGN 1\_1\_54 -runat\_26102004\_100224\_2000 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |             |        |    |                     |                    |
|------------|-------|-------------|--------|----|---------------------|--------------------|
| Result No. | Score | Query Match | Length | ID | Description         |                    |
| 1          | 430   | 45.3        | 642    | 4  | US-09-328-352-1924  | Sequence 1924, Ap  |
| 2          | 429.5 | 45.2        | 435    | 4  | US-09-489-039A-2340 | Sequence 2340, Ap  |
| 3          | 414   | 43.6        | 34063  | 3  | US-09-453-702B-96   | Sequence 96, Appl  |
| 4          | 413   | 43.5        | 48908  | 3  | US-09-453-702B-137  | Sequence 137, App  |
| 5          | 380.5 | 40.1        | 537    | 4  | US-09-543-681A-294  | Sequence 294, App  |
| 6          | 170.5 | 17.9        | 527    | 6  | 5171680-4           | Patent No. 5171680 |
| 7          | 166.5 | 17.5        | 600    | 6  | 5290690-4           | Patent No. 5290690 |
| 8          | 166.5 | 17.5        | 2523   | 1  | US-08-350-884-85    | Sequence 85, Appl  |
| 9          | 166.5 | 17.5        | 2523   | 1  | US-08-440-548-85    | Sequence 85, Appl  |
| 10         | 166.5 | 17.5        | 2523   | 1  | US-08-709-173-85    | Sequence 85, Appl  |
| 11         | 166.5 | 17.5        | 2523   | 2  | US-08-709-177-85    | Sequence 85, Appl  |
| 12         | 165.5 | 17.4        | 874    | 3  | US-09-126-109-3     | Sequence 3, Appli  |

|      |       |      |         |   |                     |                    |
|------|-------|------|---------|---|---------------------|--------------------|
| 13   | 165.5 | 17.4 | 1858    | 1 | US-08-668-381A-6    | Sequence 6, Appli  |
| 14   | 165.5 | 17.4 | 3075    | 1 | US-07-910-760-11    | Sequence 11, Appl  |
| 15   | 165.5 | 17.4 | 3075    | 1 | US-08-440-519-11    | Sequence 11, Appl  |
| 16   | 165.5 | 17.4 | 3075    | 3 | US-08-440-549-11    | Sequence 11, Appl  |
| 17   | 165.5 | 17.4 | 3297    | 4 | US-09-881-654-3     | Sequence 3, Appli  |
| 18   | 165.5 | 17.4 | 3297    | 4 | US-10-637-323-3     | Sequence 3, Appli  |
| 19   | 163.5 | 17.2 | 491     | 6 | 5252476-6           | Patent No. 5252476 |
| 20   | 163.5 | 17.2 | 617     | 6 | 5196335-2           | Patent No. 5196335 |
| 21   | 162.5 | 17.1 | 801     | 3 | US-09-291-562-1     | Sequence 1, Appli  |
| 22   | 161.5 | 17.0 | 720     | 4 | US-09-439-813-1     | Sequence 1, Appli  |
| 23   | 161.5 | 17.0 | 4403765 | 3 | US-09-103-840A-2    | Sequence 2, Appli  |
| 24   | 161.5 | 17.0 | 4411529 | 3 | US-09-103-840A-1    | Sequence 2, Appli  |
| 25   | 160   | 16.8 | 1186    | 1 | US-08-368-236-2     | Sequence 2, Appli  |
| 26   | 149.5 | 15.7 | 1080    | 1 | US-08-225-757B-1    | Sequence 1, Appli  |
| 27   | 149.5 | 15.7 | 1080    | 2 | US-08-722-050-1     | Sequence 1, Appli  |
| 28   | 149.5 | 15.7 | 1080    | 4 | US-09-883-985-1     | Sequence 1, Appli  |
| 29   | 147   | 15.5 | 480     | 4 | US-09-248-796A-6301 | Sequence 6301, Ap  |
| 30   | 138.5 | 14.6 | 515     | 4 | US-09-513-999C-1212 | Sequence 1212, Ap  |
| 31   | 117.5 | 12.4 | 747     | 3 | US-08-998-416-217   | Sequence 217, App  |
| 32   | 116.5 | 12.3 | 231     | 4 | US-09-248-796A-6300 | Sequence 6300, App |
| 33   | 95    | 10.0 | 666     | 6 | 5472691-4           | Patent No. 5472691 |
| 34   | 95    | 10.0 | 669     | 3 | US-08-556-965-1     | Sequence 1, Appli  |
| 35   | 95    | 10.0 | 723     | 6 | 5472691-7           | Patent No. 5472691 |
| 36   | 95    | 10.0 | 1396    | 6 | 5472691-1           | Patent No. 5472691 |
| 37   | 95    | 10.0 | 10079   | 2 | US-08-476-866-20    | Sequence 20, Appl  |
| C 38 | 91.5  | 9.6  | 1194    | 3 | US-08-765-907A-9    | Sequence 9, Appli  |
| C 39 | 91.5  | 9.6  | 4496    | 3 | US-08-765-907A-6    | Sequence 6, Appli  |
| C 40 | 90.5  | 9.5  | 4089    | 1 | US-07-908-245-1     | Sequence 1, Appli  |
| C 41 | 90.5  | 9.5  | 4097    | 3 | US-09-123-708-5     | Sequence 5, Appli  |
| C 42 | 90.5  | 9.5  | 4097    | 3 | US-09-123-624-5     | Sequence 5, Appli  |
| C 43 | 89.5  | 9.4  | 2159    | 3 | US-08-286-4370A-7   | Sequence 7, Appli  |
| C 44 | 88.5  | 9.3  | 2572    | 3 | US-08-804-439A-9    | Sequence 9, Appli  |
| C 45 | 88.5  | 9.3  | 2572    | 3 | US-08-720-229-9     | Sequence 9, Appli  |

ALIGNMENTS

RESULT 1  
US-09-328-352-1924  
; Sequence 1924, Application US/09328352  
; Patent No. 5562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 1924  
; LENGTH: 642  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-1924

| Alignment Scores:      |          | SUMMARIES     |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 2,61e-40 | Length:       | 642 |
| Score:                 | 430.00   | Matches:      | 89  |
| Percent Similarity:    | 64.50%   | Conservative: | 20  |
| Best Local Similarity: | 52.66%   | Mismatches:   | 56  |
| Query Match:           | 45.26%   | Indels:       | 4   |
| DB:                    | 4        | Gaps:         | 2   |

US-10-009-916A-1 (1-180) x US-09-328-352-1924 (1-642)

|    |     |  |     |
|----|-----|--|-----|
| Qy | 13  | ThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerValThrSer--- | 31  |
| Db | 133 | ACGGTAAGTAAGACTGTCTCTGATACAGCCATCCACTTACCACATGCATCAAAAAA     | 192 |
| Qy | 32  | -----GluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr     | 49  |
| Db | 193 | GTGGTGGATGTTAATGAAGTGACAGCCCAATGGTAGTATAAAAAATGGTACATTAGC    | 252 |

QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGly 69  
Db 253 TTTCAAGATAGTGATAAGGCTTAATCATTTACCCCTGCTTTAGCAGATTTTACCAAGTGGT 312  
QY 70 GluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHis 89  
Db 313 ACTCGCGGTTTCACATTCATGAATAATCCATCTTGTCGCCCTGCTGTTAAAGATGAAAA 372  
QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLysHisGlu 109  
Db 373 CACGGCGCTGCTTACTGCTGTAGCCACTATPACCCCTAACCA---GCTCCACATCAT 429  
QY 110 GlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129  
Db 430 GGTACGCCCACTACCGCGGCAATTTAGCGGATTTTACCAGCATTTAGTTGTCACATACTGGT 489  
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGluIleLysGlyArgThr 149  
Db 490 GTAGCTACAACTGCTGTTATTCACCACTTTAAACCTTGCTGACATTCGAAGTCGCGCT 549  
QY 150 ValMetIleHisAlaGlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGly 169  
Db 550 ATCATGATCCATGCTGCTGGTGATACACTATTTCAGATTCCTTTGCTCTAGTGTCGTC 609  
QY 170 GlyAlaArgIleAlaCysGlyValIle 178  
Db 610 GGAGCGCGAATTGCATGTGTGTAATC 636

RESULT 2  
US-09-489-039A-2340  
; Sequence 2340, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 2340  
; LENGTH: 435  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-2340

Alignment Scores:  
Pred. No.: 1.69e-40 Length: 435  
Score: 429.50 Matches: 81  
Percent Similarity: 70.92% Conservative: 19  
Best Local Similarity: 57.45% Mismatches: 39  
Query Match: 45.21% Indels: 2  
DB: 4 Gaps: 1

US-10-009-916A-1 (1-180) x US-09-489-039A-2340 (1-435)

QY 40 GlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnI 59  
Db 9 GCGTCGGTCAGACCATCGGAGCTCAAAATCACCAACCCGACCGCGGAGTCGAGTT 68  
QY 59 eLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyG 79  
Db 69 CGCCGCCCACTCTGCGGGCGCTACCGCGGTAAGCACGGGTTTCATATTCATGCCGAGG 128  
QY 79 ySerCysGlyProAlaGluHisAspGlyHisLeuThrThrAlaGlyLeuGlnAlaHisGlyHi 99  
Db 129 CAGCTGCCAGCGCGGCGATGAAGAAGGTAAAGCCGTGCCCGCGCGCGCGCGGAC 188  
QY 99 sTyrAspProAspLysThrGlyLysHisGlyGlyProLeuGlyAsnGlyHisLysGlyAs 119  
Db 189 TTACGATCCGACGATACCGGCAACACAGAGGCGGTTGGGGCGCGGCATCTTGGCGA 248

QY 119 pLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProAr 139  
Db 249 CTGCCCCCTGCTGCTGGTCAACGATCGGGCGGTAGCCGACCCGAGTATTATTCTCGCG 308  
QY 139 gLeu---ThrValLysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAspAs 158  
Db 309 CCTGAAAAACGCTGCGGAGGTGAAGGAAAGCGCTGATGTTCCACGTTAGGCGGGGATAA 368  
QY 158 nTyrSerAspLysProLeuProLeuGlyGlyGlyGlyAlaArgIleAlaCysGlyValIl 178  
Db 369 CATGGCCGACAGCCCGACGCTGGCGCGCGGCGGCGAACGCTTTGCCCTGCGGGGTGAT 428  
QY 178 e 178  
Db 429 T 429

RESULT 3  
US-09-453-702B-96  
; Sequence 96, Application US/09453702B  
; Patent No. 6365723  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; Burland, Valerie  
; Perna, Nicole T.  
; Plunkett, Guy  
; Welch, Rod  
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 1 South Pinckney Street  
; CITY: Madison  
; STATE: WI  
; COUNTRY: US  
; ZIP: 53701-2113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 8.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/453,702B  
; FILING DATE: 03-Dec-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/110,955  
; FILING DATE: 04-DEC-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seay, Nicholas J.  
; REGISTRATION NUMBER: 27386  
; REFERENCE/DOCKET NUMBER: 960296.95017  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (608) 251-5000  
; TELEFAX: (608) 251-9166  
; INFORMATION FOR SEQ ID NO: 96:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34063  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 96:  
US-09-453-702B-96

Alignment Scores:  
Pred. No.: 5.96e-36 Length: 34063  
Score: 414.00 Matches: 88  
Percent Similarity: 62.50% Conservative: 27  
Best Local Similarity: 47.83% Mismatches: 53  
Query Match: 43.58% Indels: 16  
DB: 3 Gaps: 5



```
US-10-009-916A-1 (1-180) x US-09-453-702B-96 (1-34063)
QY 1 MetLysIleIysLeuPheValThrSerIleValThrIleSerLeuLeuThrSerIle 20
Db 3879 ATGAATGTAAATC-----ATTGTCCTTCCCAATGCCAGCA--- 3920
QY 21 ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle 36
Db 3921 -----GCATCATCGGGATACGCGACGAGCAAGTCCCAATGAACCTTGT 3968
QY 37 AspAspAsnGlyIleGlnSerIleGlyThrValThrPheThrAspThrAspLysGly 56
Db 3969 AGTGCTGACGGAAGAAAGTACGATTTGGAATAAATACCATTCAGGAGACCCCTACGGT 4028
QY 57 LeuGlnIleIysThrAspLeuLysGlyLeuProAlaGlyClnHisGlyPheHisIleHis 76
Db 4029 CTGCTGTTACACAGCCCTTCACTCTCTGTCTGAAGGCATTCATGTTTTCATGTGCAC 4088
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db 4089 GAAAAGGAAATTCGCGCCGCGACTGAAAGACGGAACCCGGTCGCGCATTCATCGGCT 4148
QY 97 HisGlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly----AsnGly 115
Db 4149 GCGCGTCACTTTGACCGGAAAAACACCGGCAAAATCTTGGCCCTGGTCTCCGGATGGA 4208
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
Db 4209 CACCGGGGACCTTCCTGCGCTGTCGTGACCGATGACGGAAGCAACTACCCGGTC 4268
QY 136 LeuAlaProArgLeu----ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154
Db 4269 CTGCGCCGAGACTGAACATCAATAAAGAGATTAAAGGGGTTCTCTCATGCTTCATGCT 4328
QY 155 GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174
Db 4329 GCGCGTGATAACCATCATGACCATCGCGAGCCCTGGCGCGTGTGTGTGAGAAATGGCC 4388
QY 175 CysGlyValIle 178
Db 4389 TCGCGGCATCAT 4400

RESULT 4
US-09-453-702B-137
; Sequence 137, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-Dec-1998
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48908
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 137:
US-09-453-702B-137

Alignment Scores:
Pred. No.: 1.32e-35 Length: 48908
Score: 413.00 Matches: 88
Percent Similarity: 62.50% Conservative: 27
Best Local Similarity: 47.83% Mismatches: 53
Query Match: 43.47% Indels: 16
DB: 3 Gaps: 5

US-10-009-916A-1 (1-180) x US-09-453-702B-137 (1-48908)
QY 1 MetLysIleIysLeuPheValThrSerIleValThrIleSerLeuLeuThrSerIle 20
Db 6812 ATGAATGTAAATC-----ATTGTCCTTCCCAATGCCAGCA--- 6853
QY 21 ThrSerValValLeuAlaCysSerValThrSerGlu-----ValHisMetIle 36
Db 6854 -----GCATCATCGGGATACGCGACGAGCAAGTCCCAATGAACCTTGT 6901
QY 37 AspAspAsnGlyIleGlnSerIleGlyThrValThrPheThrAspThrAspLysGly 56
Db 6902 AGTGCTGACGGAAGAAAGTACGATTTGGAATAAATACCATTCAGGAGACCCCTACGGT 6961
QY 57 LeuGlnIleIysThrAspLeuLysGlyLeuProAlaGlyClnHisGlyPheHisIleHis 76
Db 6962 CTGCTGTTACACAGCCCTTCACTCTCTGTCTGAAGGCATTCATGTTTTCATGTGCAC 7021
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db 7022 GAAAAGGAAATTCGCGCCGCGACTGAAAGACGGAACCCGGTCGCGCATTCATCGGCT 7081
QY 97 HisGlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly----AsnGly 115
Db 7082 GCGCGTCACTTTGACCGGAAAAACACCGGCAAAATCTTGGCCCTGGTCTCCGGATGGA 7141
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
Db 7142 CACCTGGGACCTCCCTGCGCTGTCGTGACCGATGACGGAAGCAACTACCCGGTC 7201
QY 136 LeuAlaProArgLeu----ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154
Db 7202 CTGCGCCGAGACTGAACATCAATAAAGAGATTAAAGGGGTTCTCTCATGCTTCATGCT 7261
QY 155 GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174
Db 7262 GCGCGTGATAACCATCATGACCATCGCGAGCCCTGGCGCGTGTGTGTGAGAAATGGCC 7321
QY 175 CysGlyValIle 178
Db 7322 TCGCGGCATCAT 7333

RESULT 5
US-09-543-681A-294
; Sequence 294, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 294
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-294

Alignment Scores:
Pred. No.: 9,86e-35 Length: 537
Score: 380.50 Matches: 85
Percent Similarity: 57.65% Conservative: 13
Best Local Similarity: 50.00% Mismatches: 67
Query Match: 40.05% Indels: 5
DB: 4 Gaps: 3

US-10-009-916A-1 (1-180) x US-09-543-681A-294 (1-537)

QY 14 IleSerLeuLeuThrSer-----IleThrSerValValLeuAlaCysSerValThr 30
DB 22 ATACCTTTGTTACTCTCTGGACTACTCTTTACATCGGTTGCCAGTCCGGCGAGCTTAGAT 81
QY 31 SerGluValHisMetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThrPhe 50
DB 82 GTTACGTTAAAGAAGCCCTTACTACAGGGCGGGTGATGATTTGGTTGTTCTCACTATT 141
QY 51 ThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGlu 70
DB 142 ACCGAAACCGATTATGGCTGTATTATTACGCCAAACTTACTGGGTAAACCGGGTGT 201
QY 71 HisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeu 90
DB 202 CATGGTTTTCATATCATCTGCTAATGGCTCTTGTGAGCCAGATATGAAAGATGGAACCT 261
QY 91 ThrAlaGlyLeuGlnAlaHisGlyHisIleLysThrAspProAspLysThrGlyLysHisGly 110
DB 262 GTTCAGCATTTAAAGCGGGTGGTCATCTTGTATCTGAAATAAAGGTGTTCACTTAGGG 321
QY 111 ProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129
DB 322 CTTTATACAAAGAAGGCACTTATAGTGATCTACCTGTTGGTTGGTCCCAATAGTAAAGT 381
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThr---ValLysGluIleLysGlyArg 148
DB 382 GACGCTGATTATGCGGTTCTTGTCTCTAGACTGACTAAGCTTGATCAAAATTAAGATAAA 441
QY 149 ThrValMetIleHisAlaGlyAspAsnTyrSerAspLysProLeuProLeuGlyGly 168
DB 442 GCCTTAATGGTTACGTTGGGGAGATAACTACTCAGATAACCCAGAGCGCTTGGCGGT 501
QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
DB 502 GGTGGTGCTAGATGGCATGTGGTGTGATT 531

RESULT 6
5171680-4
;Patent No. 5171680
;APPLICANT: MULLENBACH, GUY T.;HALLEWELL, ROBERT A.;VALEZUELA,
;PABLO
;TITLE OF INVENTION: SUPEROXIDE DISMUTASE ANALOGS HAVING NOVEL
;BINDING PROPERTIES
;NUMBER OF SEQUENCES: 15
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/561,442
;FILING DATE: 01-AUG-1990
;SEQ ID NO:4
;LENGTH: 527

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 294
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-294

Alignment Scores:
Pred. No.: 9,86e-35 Length: 537
Score: 380.50 Matches: 85
Percent Similarity: 57.65% Conservative: 13
Best Local Similarity: 50.00% Mismatches: 67
Query Match: 40.05% Indels: 5
DB: 4 Gaps: 3

US-10-009-916A-1 (1-180) x US-09-543-681A-294 (1-537)

QY 14 IleSerLeuLeuThrSer-----IleThrSerValValLeuAlaCysSerValThr 30
DB 22 ATACCTTTGTTACTCTCTGGACTACTCTTTACATCGGTTGCCAGTCCGGCGAGCTTAGAT 81
QY 31 SerGluValHisMetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThrPhe 50
DB 82 GTTACGTTAAAGAAGCCCTTACTACAGGGCGGGTGATGATTTGGTTGTTCTCACTATT 141
QY 51 ThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGlu 70
DB 142 ACCGAAACCGATTATGGCTGTATTATTACGCCAAACTTACTGGGTAAACCGGGTGT 201
QY 71 HisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeu 90
DB 202 CATGGTTTTCATATCATCTGCTAATGGCTCTTGTGAGCCAGATATGAAAGATGGAACCT 261
QY 91 ThrAlaGlyLeuGlnAlaHisGlyHisIleLysThrAspProAspLysThrGlyLysHisGly 110
DB 262 GTTCAGCATTTAAAGCGGGTGGTCATCTTGTATCTGAAATAAAGGTGTTCACTTAGGG 321
QY 111 ProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129
DB 322 CTTTATACAAAGAAGGCACTTATAGTGATCTACCTGTTGGTTGGTCCCAATAGTAAAGT 381
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThr---ValLysGluIleLysGlyArg 148
DB 382 GACGCTGATTATGCGGTTCTTGTCTCTAGACTGACTAAGCTTGATCAAAATTAAGATAAA 441
QY 149 ThrValMetIleHisAlaGlyAspAsnTyrSerAspLysProLeuProLeuGlyGly 168
DB 442 GCCTTAATGGTTACGTTGGGGAGATAACTACTCAGATAACCCAGAGCGCTTGGCGGT 501
QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
DB 502 GGTGGTGCTAGATGGCATGTGGTGTGATT 531

RESULT 7
5290690-4
;Patent No. 5290690
;APPLICANT: MRABET, NADIR;LASTERS, IGNACE;STANSSENS, PATRICK
;MATHYSSENS, GASTON;WODAK, SHOSHANA;QUAX, WILHELMUS J.
;TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
;STABILITY OF PROTEINS
;NUMBER OF SEQUENCES: 22
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/398,706
;FILING DATE: 25-AUG-1989
;SEQ ID NO:4
;LENGTH: 600
5290690-4

Alignment Scores:
Pred. No.: 4,64e-10 Length: 600
Score: 166.50 Matches: 52
Percent Similarity: 47.85% Conservative: 26
Best Local Similarity: 31.90% Mismatches: 56
Query Match: 17.53% Indels: 29
DB: 6 Gaps: 9

US-10-009-916A-1 (1-180) x 5290690-4 (1-600)
```

```
5171680-4

Alignment Scores:
Pred. No.: 1,33e-10 Length: 527
Score: 170.50 Matches: 54
Percent Similarity: 45.98% Conservative: 26
Best Local Similarity: 31.03% Mismatches: 63
Query Match: 17.95% Indels: 31
DB: 6 Gaps: 10

US-10-009-916A-1 (1-180) x 5171680-4 (1-527)

QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
DB 8 ACAAGGCTGTTGCTGTTTGAAGGTTGACGGCGCGTTCAA-----GGTATTATTAAAC 61
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
DB 62 TTCGAGCAGAGAGAAAGTAATCGACCACTGTAAGAGGTGTGGGGAAGCATTTAAGGACTGCT 121
QY 68 AlaGlyGluHisGlyPheHisIleHisGlu-----GlyGlySerCysGly 82
DB 122 GAAGCCTCGCATGGATTCCATGTTTCATGAGTTTGGAGATAATACAGCAGGCTGTACCAGT 181
QY 83 ProAlaGluHisAspGlyHisLeu-----ThrAlaGlyLeuGlnAla 96
DB 182 GAAGCCTCGCATGGATTCCATGTTTTCATGAGTTTGGAGATAATACAGCAGGCTGTACCAGT 241
QY 97 HisGly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
DB 242 GCAGGTCTCTCACTTTAATCTCT---CTCTCGAGAAACACACGGTGGGCCCAAGGATGAAGAG 298
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
DB 299 AGGCATGTTGGAGACTTGGCNAATGTGACCCCGGACAAAGATGGTGTGCCCGCATGTATCG 358
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151
DB 359 ATTGAAGATTCTGTGATATCACTCTCAGGAGACCACTTCCATCATTTGGCGCACACTAGTG 418
QY 152 IleHis-----AlaGlyGlyAspAsnTyrSerAspLysProLeu 164
DB 419 GTCCATGAAAAAGCAGATGACTTGGGCAAGAGTGGAAATGAAGAAAGTACAAAG----- 472
QY 165 ProLeuGlyGlyGlyValaArgIleAlaCysGlyValIle 178
DB 473 ---ACAGAAACCGTGGATCCCGTTTGGCTTGTGTGTAATT 511

RESULT 7
5290690-4
;Patent No. 5290690
;APPLICANT: MRABET, NADIR;LASTERS, IGNACE;STANSSENS, PATRICK
;MATHYSSENS, GASTON;WODAK, SHOSHANA;QUAX, WILHELMUS J.
;TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
;STABILITY OF PROTEINS
;NUMBER OF SEQUENCES: 22
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/398,706
;FILING DATE: 25-AUG-1989
;SEQ ID NO:4
;LENGTH: 600
5290690-4

Alignment Scores:
Pred. No.: 4,64e-10 Length: 600
Score: 166.50 Matches: 52
Percent Similarity: 47.85% Conservative: 26
Best Local Similarity: 31.90% Mismatches: 56
Query Match: 17.53% Indels: 29
DB: 6 Gaps: 9

US-10-009-916A-1 (1-180) x 5290690-4 (1-600)
```



```

, ZIP: 94304-1018
,
, COMPUTER READABLE FORM:
,
, MEDIUM TYPE: Floppy disk
,
, COMPUTER: IBM PC compatible
,
, OPERATING SYSTEM: PC-DOS/MS-DOS
,
, SOFTWARE: PatentIn Release #1.0, Version #1.25
,
, CURRENT APPLICATION DATA:
,
, APPLICATION NUMBER: US/08/440,548
,
, FILING DATE: 12-MAY-1995
,
, CLASSIFICATION: 435
,
, PRIOR APPLICATION DATA:
,
, APPLICATION NUMBER: US 07/680,296
,
, FILING DATE: 04-APR-1991
,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: CIOTTI, THOMAS E.
,
, REGISTRATION NUMBER: 21,013
,
, REFERENCE/DOCKET NUMBER: 22300-20100.20
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: (415) 813-5600
,
, TELEFAX: (415) 494-0792
,
, TELEX: 706141
,
, INFORMATION FOR SEQ ID NO: 85:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 2523 base pairs
,
, TYPE: nucleic acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, FEATURE:
,
, NAME/KEY: CDS
,
, LOCATION: 1..2523
,
, US-08-440-548-85

```

|                        |           |                  |
|------------------------|-----------|------------------|
| Alignment Scores:      |           |                  |
| Pred. No.:             | 3, 78e-09 | 2523             |
| Score:                 | 166.50    | 52               |
| Percent Similarity:    | 47.85%    | Conservative: 26 |
| Best Local Similarity: | 31.90%    | Mismatches: 56   |
| Query Match:           | 17.53%    | Indels: 29       |
| DB:                    | 1         | Gaps: 9          |

US-10-009-916A-1 (1-180) x US-08-440-548-85 (1-2523)

|     |   |     |
|-----|---|-----|
| 30  | ThrSerGluValHisMetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThr    | 49  |
| 7   | ACAAACCCCTGTTTGGCTTTGAAGGGTGCAGGCCAGTCAA-----CGTATTATTAAAC      | 60  |
| 50  | PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro     | 67  |
| 61  | TTTCGAGCAGACAGAAAGTAATGGACCGTAGGTGGGGAAGCATTAAGGACTGACT         | 120 |
| 68  | AlaGlyGluHisGlyPheHisIleGlnGlyGlySerCysGlyProAlaGluHisAsp       | 87  |
| 121 | GAAGGCCCTCATGGATTCCATGTTTCATGAGTTTCGAGAT-----                   | 159 |
| 88  | GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTryAspProAspLysThrGly    | 106 |
| 160 | -----ATACAGCAGCGGTGCACAGCTCCAGGTCTCTCACTTTAATCTCT-----CTATCCAGA | 210 |
| 107 | LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProAsnGluVal       | 125 |
| 211 | AAACACGGTGGGCCCAAGGATCAAGACAGAGGCGATGTTGGAGACTTGGCAATGTGACTGCT  | 270 |
| 126 | LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu---    | 144 |
| 271 | GACAAAGATGGTGTGGCCGATGTGCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC     | 330 |
| 145 | -----IleLysGlyArgThrValMetIleHis-----AlaGly                     | 155 |
| 331 | CATTGCATCATTTGGCCGCGACACTGGTGGTCCATGAAAAAGCAGATGACTTGGCAAAAGT   | 390 |
| 156 | GlyAspAsnTrySerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys       | 175 |
| 391 | GGAAATTGAAGAAAGTACAAAG-----ACAGGAACCGCTGGAAGTCGTTGGCTGT         | 441 |

```
Db 121 GAAGGCTGCATGATTCATGCTTGGAGAT----- 159
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106
Db 160 -----AATACAGAGGCTGTACCGAGTCCAGTCTCACTTTAAATCT---CTATCCAGA 210
QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125
Db 211 AAACACGGTGGCCCAAGAGTGAAGAGGAGCATGTTGGAGACTTGGCAATGTGACTGCT 270
QY 126 LysAlaAspGlyLeuAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu---- 144
Db 271 GACAAGATGGTGGCCGACACTGGTGGTCCATGAAAGAGCAGATGACTTGGGCAAAAGT 330
QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 331 CATTCGATCATTTGGCCGACACTGGTGGTCCATGAAAGAGCAGATGACTTGGGCAAAAGT 390
QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaAlaCys 175
Db 391 GGAATGAAGAAAGTACAAAG-----ACAGGAAACGCTGGAAGTCTTTGGCTTGT 441
QY 176 GlyValIle 178
Db 442 GGTGTAAT 450

RESULT 11
US-08-709-177-85
; Sequence 85, Application US/08709177
; Patent No. 5885799
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI LIM
; APPLICANT: KUO, GEORGE
; TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,177
; FILING DATE: 06-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/680,296
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CIOTTI, THOMAS E.
; REGISTRATION NUMBER: 21,013
; REFERENCE/DOCKET NUMBER: 22300-20100.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2523
; US-08-709-177-85
```

```
Alignment Scores:
Pred. No.: 3 78e-09 Length: 2523
Score: 166.50 Matches: 52
Percent Similarity: 47.8% Conservative: 26
Best Local Similarity: 31.90% Mismatches: 56
Query Match: 17.53% Indels: 9
DB: 2 Gaps: 9

US-10-009-916A-1 (1-180) x US-08-709-177-85 (1-2523)
QY 30 ThrSerGluValHisMetIleAspAsnGlyLeuLysGlnSerIleGlyThrValThr 49
Db 7 ACAACCCCTGTTGGCTTTTGAAGGCTGACGCCGCCAGTTCAA-----GGTATTATTAAC 60
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 61 TTTCAGCAGAGAAGAAAGTAATGGACACAGTGAAGGCTGGGGAAGCATTTAAAGGACTGACT 120
QY 68 AlaGlyGluHisGlyPheHisIleHisGlyGlySerCysGlyProAlaGluHisAsp 87
Db 121 GAAGCCTGCATGGATTCCATGTTTCATGAGTTGGAGAT----- 159
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106
Db 160 -----AATACAGAGGCTGTACCGAGTCCAGTCTCACTTTAAATCT---CTATCCAGA 210
QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125
Db 211 AAACACGGTGGCCCAAGAGTGAAGAGGAGCATGTTGGAGACTTGGGCAATGTGACTGCT 270
QY 126 LysAlaAspGlyLeuAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144
Db 271 GACAAGATGGTGGCCGACACTGGTGGTCCATGAAAGAGCAGATGACTTGGGCAAGGT 330
QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 331 CATTCGATCATTTGGCCGACACTGGTGGTCCATGAAAGAGCAGATGACTTGGGCAAGGT 390
QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaAlaCys 175
Db 391 GGAATGAAGAAAGTACAAAG-----ACAGGAAACGCTGGAAGTCTTTGGCTTGT 441
QY 176 GlyValIle 178
Db 442 GGTGTAAT 450

RESULT 12
US-09-126-109-3
; Sequence 3, Application US/09126109
; Patent No. 6171856
; GENERAL INFORMATION:
; APPLICANT: Thigpen, Anice
; APPLICANT: Hohmeier, Hans-Ewald
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Unger, Roger H.
; APPLICANT: Shimabukuro, Michio
; APPLICANT: Chen, Guaxun
; APPLICANT: Rhodes, Christopher J.
; APPLICANT: Hugl, Sigrun R.
; APPLICANT: Cousin, Sharon
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING
; TO NO-MEDIATED CYTOTOXICITY
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,109  
FILING DATE: 30-JUL-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/055,092  
FILING DATE: 30-JUL-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US Unknown  
FILING DATE: 03-MAR-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: McMillian, Nabeela R.  
REGISTRATION NUMBER: P-43,363  
REFERENCE/DOCKET NUMBER: UTSID:560  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 874 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-126-109-3

Alignment Scores:  
Pred. No.: 1,05e-09 Length: 874  
Score: 165.50 Matches: 52  
Percent Similarity: 47.24% Conservative: 25  
Best Local Similarity: 31.90% Mismatches: 57  
Query Match: 17.42% Indels: 29  
DB: 3 Gaps: 9

US-10-009-916A-1 (1-180) x US-09-126-109-3 (1-874)

QY 30 ThrSerGluValHisMetIleAspAsnGlyLeuGlnSerIleGlyThrValThr 49  
Db 71 ACAGAGCGGTGTCGTCTGAGGGCGACGGCCAGTCGAG-----GGCATCATCAAT 124  
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67  
Db 125 TTCGAGCAGAGAGAAAGTAATGACCACTGAAGGTGTGGGGAAGCATTAAGGACTGACT 184  
QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87  
Db 185 GAAGCGCTGCATGGATTCCATGTCATGATTTGGAGAT----- 223  
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106  
Db 224 -----AATACAGCAGCGGTGTACCACTGTCAGTCTCTCACTTAACTCT---CTATCCAGA 274  
QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125  
Db 275 AAACACGGTGGGCCCAAGGATCAAGAGGCGATGTTGGAGATCTGGGCAATGTGACTGCT 334  
QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144  
Db 335 GACAAAGATGTGTGGCGCATGTGCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 394  
QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155  
Db 395 CATTCATCATTTGGCCGACACTGTGTGTCATGAAAAGCAGATGACTTGGGCAGAGGT 454  
QY 156 GlyAspAsnTyrSerAspLysProLeuGlyGlyGlyAlaArgIleAlaCys 175  
Db 455 GGAATGAAAGAAAGTACAAAG-----ACAGGAAACCGTGGAAAGTCGTTGGCTTGT 505  
QY 176 GlyValIle 178  
|||||

Db 506 GGTGTAATT 514  
RESULT 13  
US-08-668-381A-6  
; Sequence 6, Application US/08668381A  
; Patent No. 5780024  
; GENERAL INFORMATION:  
; APPLICANT: Brown, Robert H.  
; APPLICANT: Fishman, Paul S.  
; APPLICANT: Francis, Jonathan W.  
; APPLICANT: Hosler, Betsy A.  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN  
; TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN  
; CORRESPONDENCE ADDRESSES:  
; NUMBER OF SEQUENCES: 6  
; ADDRESS: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/668,381A  
; FILING DATE: 21-JUN-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/000,473  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,164  
; REFERENCE/DOCKET NUMBER: 00786/269001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1858 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-668-381A-6  
Alignment Scores:  
Pred. No.: 3.15e-09 Length: 1858  
Score: 165.50 Matches: 52  
Percent Similarity: 47.24% Conservative: 25  
Best Local Similarity: 31.90% Mismatches: 57  
Query Match: 17.42% Indels: 29  
DB: 1 Gaps: 9  
US-10-009-916A-1 (1-180) x US-08-668-381A-6 (1-1858)  
QY 30 ThrSerGluValHisMetIleAspAsnGlyLeuGlnSerIleGlyThrValThr 49  
Db 8 ACAGAGCGGTGTCGTCTGAGGGCGACGGCCAGTCGAG-----GGCATCATCAAT 61  
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67  
Db 62 TTCGAGCAGAGAGAAAGTAATGACCACTGAAGGTGTGGGGAAGCATTAAGGACTGACT 121  
QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87  
Db 122 GAAGCGCTGCATGGATTCCATGTCATGATTTGGAGAT----- 160  
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106

Best Local Similarity: 31.90% Mismatches: 57  
 Query Match: 17.42% Indels: 29  
 DB: 1 Gaps: 9

US-10-009-916A-1 (1-180) x US-07-910-760-11 (1-3075)

QY 30 ThrSerGluValHisMetileAspAspAspGlyLeuGlnSerileGlyThrValThr 49  
 DB 7 ACAAGAGCTGTTTGTGTTTGAAGGTGACGGCCAGTTCNA-----GGTATTATTAAC 60  
 QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67  
 DB 61 TTGACAGCAGAGAAAGTAATGACACAGTCAAGGTGTGGGAAAGCATTAAAGAGCTGACT 120  
 QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87  
 DB 121 GAAGGCTGCATGATTCATGTTTCATGAGTTTGAGAT----- 159  
 QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106  
 DB 160 -----AATACAGCAGGCTGTACAGTGCAGGTCTCCTCACTTTAACTCT---CTATCCAGA 210  
 QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125  
 DB 211 AACACGGTGGCCCAAGGATGAAGAGGATGTTGGAGACTTGGCAATGTGCTGCT 270  
 QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144  
 DB 271 GACAAAGATGTTGTGGCCGATGTTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 330  
 QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155  
 DB 331 CATTTGCATCATTTGGCCGACACTGTTGTTCTCCGAAAAAGCAGATGACTTGGCAAGGT 390  
 QY 156 GlyAspAsnTyrSerAspLysProLeuGlyGlyGlyGlyAlaArgIleAlaCys 175  
 DB 391 GGAATGAAGAAAGTACAAAG-----ACAGAAACGCTGGAAGTCTGTTGGCTGT 441

176 GlyValIle 178  
 442 GGTGTAAT 450

RESULT 15  
 US-08-440-519-11  
 ; Sequence 11, Application US/08440519  
 ; Patent No. 5712087  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Houghton, Michael  
 ; APPLICANT: Choo, Qui-Lim  
 ; APPLICANT: Kuo, George  
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
 ; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Chiron Corporation  
 ; STREET: P.O. Box 8097 (Int. Prop. R-440)  
 ; CITY: Emeryville  
 ; STATE: CA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 94662-8097  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/910,760  
 ; FILING DATE: 07-JUL-1992  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Blackburn Esq., Robert P.  
 ; REGISTRATION NUMBER: 30,447  
 ; REFERENCE/DOCKET NUMBER: 0101.002  
 ; TELEPHONE: (510) 601-2702  
 ; TELEFAX: (510) 655-3542  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3075 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..3063  
 ; US-07-910-760-11

DB 161 -----AATACAGCAGGCTGTACAGTGCAGGTCTCACTTAATCT---CTATCCAGA 211  
 QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125  
 DB 212 AACACGGTGGCCCAAGGATGAAGAGGATGTTGGAGACTTGGCAATGTGACTGCT 271  
 QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144  
 DB 272 GACAAAGATGTTGTGGCCGATGTTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 331  
 QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155  
 DB 332 CATTTGCATCATTTGGCCGACACTGTTGTTCTCCGAAAAAGCAGATGACTTGGCAAGGT 391  
 QY 156 GlyAspAsnTyrSerAspLysProLeuGlyGlyGlyGlyAlaArgIleAlaCys 175  
 DB 392 GGAATGAAGAAAGTACAAAG-----ACAGAAACGCTGGAAGTCTGTTGGCTGT 442

176 GlyValIle 178  
 443 GGTGTAAT 451

RESULT 14  
 US-07-910-760-11  
 ; Sequence 11, Application US/07910760  
 ; Patent No. 5683864  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Houghton, Michael  
 ; APPLICANT: Choo, Qui-Lim  
 ; APPLICANT: Kuo, George  
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
 ; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Chiron Corporation  
 ; STREET: P.O. Box 8097 (Int. Prop. R-440)  
 ; CITY: Emeryville  
 ; STATE: CA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 94662-8097  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/910,760  
 ; FILING DATE: 07-JUL-1992  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Blackburn Esq., Robert P.  
 ; REGISTRATION NUMBER: 30,447  
 ; REFERENCE/DOCKET NUMBER: 0101.002  
 ; TELEPHONE: (510) 601-2702  
 ; TELEFAX: (510) 655-3542  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3075 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..3063  
 ; US-07-910-760-11

Alignment Scores: 6.57e-09 Length: 3075  
 Pred. No.: 165.50 Matches: 52  
 Score: 47.24% Conservative: 25  
 Percent Similarity:

```
/ NAME: Blackburn Esq., Robert P.
/ REGISTRATION NUMBER: 30,447
/ REFERENCE/DOCKET NUMBER: 0101.002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (510) 601-2702
/ TELEFAX: (510) 655-3542
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3075 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..3063
US-08-440-519-11

Alignment Scores:
Pred. No.: 6.57e-09 Length: 3075
Score: 165.50 Matches: 52
Percent Similarity: 47.24% Conservative: 25
Best Local Similarity: 31.90% Mismatches: 57
Query Match: 17.42% Indels: 29
DB: 9 Gaps: 1

US-10-009-916A-1 (1-180) x US-08-440-519-11 (1-3075)
QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 7 ACAAGGCTGTTGTTGTTGAAGGCTGACGCCCGAGTTCAA-----GGTATTATTAAC 60
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 61 TTCGACGAGAGGAAGTAATGACCCAGTGAAGCTGTGGGAGACATTAAGGACTGACT 120
QY 68 AlaGlyGluHisGlyPheHisIleHisGlySerCysGlyProAlaGluHisAsp 87
Db 121 GAAGGCTGCATGGATTCCATGTTTCATGATTTGGAGAT-----159
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106
Db 160 -----AATACAGCAGGCTGTACAGTGCAGGCTCTCACTTTAATCCT---CTATCCAGA 210
QY 107 LysHisGluGlyProLeuGly---AsnGlyHisGlyAspLeuProArgLeuValVal 125
Db 211 AAACACGGTGGCCCAAGAGTGAAGAGAGGACATGTTGGAGACTTGGGCAATGTGACTGCT 270
QY 126 LysAlaAspGlyIleAlaLysGluThrLeuAlaProArgLeuThrValLysGlu--- 144
Db 271 GACAAAGATGGTGGCCGATGTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 330
QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 331 CATTGCATCATTTGGCCGACACTGGTGTCCATGAAAGAGCAGATGACTTGGGCAAGGT 390
QY 156 GlyAspAsnTyrSerAspLysProLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 175
Db 391 GGAATGAGAAAGTACAAG-----ACAGGAAACGGTGAAGTGGTGGCTGTG 441
QY 176 GlyValIle 178
Db 442 GGTGAAT 450

RESULT 16
US-08-440-549-11
/ Sequence 11, Application US/08440549
/ Patent No. 6312889
/ GENERAL INFORMATION:
/ APPLICANT: Houghton, Michael
/ APPLICANT: Choo, Qui-Lim
/ APPLICANT: Kuo, George
/ TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)

/ TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Chiron Corporation
/ STREET: P.O. Box 8097 (Int. Prop. R-440)
/ CITY: Emeryville
/ STATE: CA
/ COUNTRY: U.S.A.
/ ZIP: 94662-8097
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/440,549
/ FILING DATE: 12-MAY-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/910,760
/ FILING DATE: 07-JUL-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Blackburn Esq., Robert P.
/ REGISTRATION NUMBER: 30,447
/ REFERENCE/DOCKET NUMBER: 0101.002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (510) 601-2702
/ TELEFAX: (510) 655-3542
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3075 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..3063
US-08-440-549-11

Alignment Scores:
Pred. No.: 6.57e-09 Length: 3075
Score: 165.50 Matches: 52
Percent Similarity: 47.24% Conservative: 25
Best Local Similarity: 31.90% Mismatches: 57
Query Match: 17.42% Indels: 29
DB: 9 Gaps: 1
```



QY 145 -----IleYsGlyArgThrValMetIleHis-----AlaGly 155  
Db 331 CATTGCATCATTGGCGCCACACTGGTGGCCATGAARAAGACAGATGACTTGGCRAAGGT 390  
QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175  
Db 391 GGAATGAAGAAAGTACAAG-----ACAGGAAACCGCTGGAAGTGGTTGGCTTGT 441  
QY 176 GlyValIle 178  
Db 442 GGTGTAATT 450  
RESULT 17  
US-09-881-654-3  
; Sequence 3, Application US/09881654  
; Patent No. 6632601  
; GENERAL INFORMATION:  
; APPLICANT: CHIEN, David Y.  
; APPLICANT: ARANGEL, Phillip  
; APPLICANT: TANDESKE, Laura  
; APPLICANT: GEORGE-NASCIEMENTO, Carlos  
; APPLICANT: COLT, Doris  
; APPLICANT: MEDINA-SELBY, Angelica  
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES  
; FILE REFERENCE: 2302-17039 / P17039.002  
; CURRENT APPLICATION NUMBER: US/09/881,654  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/212,082  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/280,811  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/280,867  
; PRIOR FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 3297  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1  
; NAME/KEY: CDS  
; LOCATION: (1)..(3297)  
US-09-881-654-3  
Alignment Scores:  
Pred. No.: 7.28e-09 Length: 3297  
Score: 165.50 Matches: 52  
Percent Similarity: 47.24% Conservative: 25  
Best Local Similarity: 31.90% Mismatches: 57  
Query Match: 17.42% Indels: 29  
DB: 4 Gaps: 9  
US-10-009-916A-1 (1-180) x US-09-881-654-3 (1-3297)  
QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49  
Db 7 ACAAAAGGCTGTGTGTTTGAAGGCGTACGCCGCCAGTTCAA-----GGTATTATTAAAC 60  
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67  
Db 61 TTCGAGCAGAGAAAGTAATGACAGGAGTGTGGGAGGAGCATTAAGAGTACTGACT 120  
QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87  
Db 121 GAAGGCTCAGTATTCATGTCATGATGTTGGAGAT-----159  
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly-----HisTyrAspProAspLysThrGly 106  
Db 160 -----AATACAGCAGGCTGTACCATGCGAGGTCTCTCACTTAATCTCT---CTATCCAGA 210  
QY 107 LysHisGluGlyProLeuGly-----AenGlyHisLysGlyAspLeuProArgLeuValVal 125

Db 211 AAACACGGTGGCCCAAGAGATGAAGAGGCAATGTTGGAGACTTGGCAATGACTGCT 270  
QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144  
Db 271 GACAAAGATGGTGGCCGATGTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 330  
QY 145 -----IleYsGlyArgThrValMetIleHis-----AlaGly 155  
Db 331 CATTGCATCATTGGCGCCACACTGGTGGTCCATGAARAAGACAGATGACTTGGGCAAGGT 390  
QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175  
Db 391 GGAATGAAGAAAGTACAAG-----ACAGGAAACCGCTGGAAGTGGTTGGCTTGT 441  
QY 176 GlyValIle 178  
Db 442 GGTGTAATT 450  
RESULT 18  
US-10-637-323-3  
; Sequence 3, Application US/10637323  
; Patent No. 6797809  
; GENERAL INFORMATION:  
; APPLICANT: CHIEN, David Y.  
; APPLICANT: ARANGEL, Phillip  
; APPLICANT: TANDESKE, Laura  
; APPLICANT: GEORGE-NASCIEMENTO, Carlos  
; APPLICANT: COLT, Doris  
; APPLICANT: MEDINA-SELBY, Angelica  
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES  
; FILE REFERENCE: 2302-17039 / P17039.002  
; CURRENT APPLICATION NUMBER: US/10/637,323  
; CURRENT FILING DATE: 2003-08-08  
; PRIOR APPLICATION NUMBER: US/09/881,654  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/212,082  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/280,811  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/280,867  
; PRIOR FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 3297  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1  
; NAME/KEY: CDS  
; LOCATION: (1)..(3297)  
US-10-637-323-3  
Alignment Scores:  
Pred. No.: 7.28e-09 Length: 3297  
Score: 165.50 Matches: 52  
Percent Similarity: 47.24% Conservative: 25  
Best Local Similarity: 31.90% Mismatches: 57  
Query Match: 17.42% Indels: 29  
DB: 4 Gaps: 9  
US-10-009-916A-1 (1-180) x US-10-637-323-3 (1-3297)  
QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49  
Db 7 ACAAAAGGCTGTGTGTTTGAAGGCGTACGCCGCCAGTTCAA-----GGTATTATTAAAC 60  
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67  
Db 61 TTCGAGCAGAGAAAGTAATGACAGGAGTGTGGGAGGAGCATTAAGAGTACTGACT 120

QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87  
Db 121 GAAGGCTGATGATTCATGTTTCATGAGTTTGGAGAT----- 159  
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106  
Db 160 -----AATACAGCAGGCTGTACAGTGCAGGCTCCTCACTTTAATCCT---CTATCCAGA 210  
QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProAlaGLeuValVal 125  
Db 211 AAACACGGTGGGCCAAAGAGTGAAGAGAGCATGTTGGAGACTTGGGCAATGTGACTGCT 270  
QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144  
Db 271 GACAAAGATGGTGGCCGATGTCTATTCAAGATTCTGTGATCTCACTCTCAGGAGAC 330  
QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155  
Db 331 CATTGCATCATTTGGCCGCACACTGGTGTCTCCATGAAAGCAGATGACTTGGGCAAGGT 390  
QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175  
Db 391 GGAATGAAGAAAGTACAAAG-----ACAGGAACGCTGGAAGTCTTTGGCTTGT 441  
QY 176 GlyValIle 178  
Db 442 GGTGTAATT 450

RESULT 19  
5252476-6  
; Patent No. 5252476  
; APPLICANT: HALLEWELL, ROBERT A.; MULLENBACH, GUY T.  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE CLONING AND  
; EXPRESSION IN MICROORGANISMS  
; NUMBER OF SEQUENCES: 15  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/222,352  
; FILING DATE: 20-JUL-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 931,920  
; FILING DATE: 14-NOV-1986  
; APPLICATION NUMBER: 609,412  
; FILING DATE: 11-MAY-1984  
; APPLICATION NUMBER: 538,607  
; FILING DATE: 03-OCT-1983  
; SEQ ID NO: 6:  
; LENGTH: 491  
5252476-6

Alignment Scores:  
Pred. No.: 7,65e-10 Length: 491  
Score: 163.50 Matches: 48  
Percent Similarity: 47.62% Conservative: 22  
Best Local Similarity: 32.65% Mismatches: 50  
Query Match: 17.21% Indels: 27  
DB: 6 Gaps: 8

US-10-009-916A-1 (1-180) x 5252476-6 (1-491)  
QY 46 GlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeu 63  
Db 19 GGCATCATCAATTTGACGACGAAGAAAGTAAATGGACCAGTGAAGGTGTGGGAAGCAT 78  
QY 64 LysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyPro 83  
Db 79 AAAGGACTGACTGAAGGCTGCATGGATTCCATGTTTCATGATTGGAGAT----- 129  
QY 84 AlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspPro 102  
Db 130 -----AATACAGCAGGCTGTACCATGTCCTCACTTCAATCCT 171  
QY 103 AspLysThrGlyLysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuPro 121  
Db 103 AspLysThrGlyLysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuPro 121

Db 172 ---CTATCCAGAAAACACGGTGGCCCAAAGATGAAGAGGCAATGTTGGAGACTTGGGC 228  
QY 122 ArgLeuValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThr 141  
Db 229 AATGTGACTGCTGACAAAGATGGTGGCCGATGTCTATTGAAGATTCTGTGATCTCA 288  
QY 142 ValLysGlu-----IleLysGlyArgThrValMetIleHis----- 153  
Db 289 CTCTCAGGAGACCATTCATTCATGGCCGCACACTGGTGTGTCATGAAAGCAGATGAC 348  
QY 154 -----AlaGlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAla 171  
Db 349 TTGGGCAAAAGTGAATGAAGAAAGTACAAAG-----ACAGGAACGCTGGCGGT 399  
QY 172 ArgIleAlaCysGlyValIle 178  
Db 400 CGTTGGCTTGTGTGTAATT 420

RESULT 20  
5196335-2  
; Patent No. 5196335  
; APPLICANT: GRONER, YORAM  
; TITLE OF INVENTION: HUMAN SUPEROXIDE DISMUTASE CDNA  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/547,827  
; FILING DATE: 2-JUL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 315,331  
; FILING DATE: 24-FEB-1989  
; APPLICATION NUMBER: 726,500  
; FILING DATE: 24-APR-1985  
; APPLICATION NUMBER: 489,786  
; FILING DATE: 29-APR-1983  
; SEQ ID NO: 2:  
; LENGTH: 617  
5196335-2

Alignment Scores:  
Pred. No.: 1.07e-09 Length: 617  
Score: 163.50 Matches: 48  
Percent Similarity: 47.62% Conservative: 22  
Best Local Similarity: 32.65% Mismatches: 50  
Query Match: 17.21% Indels: 27  
DB: 6 Gaps: 8

US-10-009-916A-1 (1-180) x 5196335-2 (1-617)  
QY 46 GlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeu 63  
Db 78 GGCATCATCAATTTGACGACGAAGAAAGTAAATGGACCAGTGAAGGTGTGGGAAGCAT 137  
QY 64 LysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyPro 83  
Db 138 AAAGGACTGACTGAAGGCTGCATGGATTCCATGTTTCATGATTGGAGAT----- 188  
QY 84 AlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspPro 102  
Db 189 -----AATACGCGAGGCTGTACCATGTCAGGTCCTCACTTTAATCCT 230  
QY 103 AspLysThrGlyLysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuPro 121  
Db 231 ---CTATCCAGAAAACACGGTGGCCCAAAGATGAAGAGGCAATGTTGGAGACTTGGGC 287  
QY 122 ArgLeuValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThr 141  
Db 288 AATGTGACTGCTGACAAAGATGGTGGCCGATGTCTATTGAAGATTCTGTGATCTCA 347  
QY 142 ValLysGlu-----IleLysGlyArgThrValMetIleHis----- 153  
Db 348 CTCTCAGGAGACCATTCATTCATTTGGCCGCACACTGGTGTGTCATGAAAGCAGATGAC 407  
QY 154 -----AlaGlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAla 171

```
Db 408 TTGGGCAAGGTGGAATGAAGAAAGTACAAAG-----ACAGAAACGCTGGAAGT 458
Qy 172 ArgileAlaCysGlyValle 178
Db 459 CGTTTGGCTTGTGGTGAATT 479

RESULT 21
US-09-291-562-1
; Sequence 1, Application US/09291562
; Patent No. 6084152
; GENERAL INFORMATION:
; APPLICANT: Sang Soo Kwak
; APPLICANT: Jae-Whune Kim
; APPLICANT: Jaeng-Soon Lee
; APPLICANT: Suk Yoon Kwon
; TITLE OF INVENTION: METHOD FOR PRODUCING TRANSGENIC CUCUMBER
; FILE REFERENCE: 118.1-US-01
; CURRENT APPLICATION NUMBER: US/09/291,562
; CURRENT FILING DATE: 1999-04-14
; EARLIER FILING DATE: 1998-04-14
; EARLIER APPLICATION NUMBER: KS 98 13205
; EARLIER FILING DATE: 1998-04-14
; EARLIER APPLICATION NUMBER: KS 98 33947
; EARLIER FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: KS 99 11848
; EARLIER FILING DATE: 1999-04-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Manihot esculenta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55)...(510)
; NAME/KEY: polyA site
; LOCATION: (781)...(801)
; NAME/KEY: polyA signal
; LOCATION: (611)...(616)
US-09-291-562-1

Alignment Scores:
Pred. No.: 2,04e-09 Length: 801
Score: 162.50 Matches: 49
Percent Similarity: 42.42% Conservative: 21
Best Local Similarity: 29.70% Mismatches: 62
Query Match: 17.11% Indels: 33
DB: 3 Gaps: 8

US-10-009-916A-1 (1-180) x US-09-291-562-1 (1-801)
Qy 29 ValThrSerGluValHisMetIleAspAsnGlyLeuLysGlnSerIleGlyThrVal 48
Db 58 GTGAAGGCTGAAGTGTCTTACCAGTAGTAGGGGGTTAGC-----GGAACAATC 108
Qy 49 ThrPheThrAspThrAspLysGly---LeuGlnIleLysThrAspLeuLysGlyLeuPro 67
Db 109 TTTCTTACCCAAAGAGGATGGTCTACCACTGTAACCTGGAAACATTTCCGGCCTTAAG 168
Qy 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGly-----Ser 80
Db 169 CCAGGCTTCATGGTTCACGCTCCATGCCCTTGGAGACACAAACGGTTGCATGTCA 228
Qy 81 CysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyr 100
Db 229 ACTGGGCA-----CACTTTAAACCTTCTGGCAAAAGATCATGGT----- 267
Qy 101 AspProAspLysThrGlyIleHisGluGlyProLeuGlyAsnGlyHisLysGlyAspLeu 120
Db 268 GCCCCTGAGATGAGATTCGTATCGTATCGTATCGTATCGTATCGTATCGTATCGTAT 324
Qy 121 ProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeu 140
```

```
Db 325 -----GATGGCACTGCTAGTTTCACAATTATTGCAAGCATATT 363
Qy 141 ThrVal-----LysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAsp 157
Db 364 CTTCTTTCTGGTCAAAATTCATATAGGAAGGCGAGTTGTTTTCATGCAGATCCTGAT 423
Qy 158 AsnTyrSer-----AspLysProLeuProLeuGlyGlyGlyAlaArgIle 173
Db 424 GATCTTGGCAGGGGAGGACATGAACCTAGTAAACACCGGAAATGCTGGTGCGAGAGTA 483
Qy 174 AlaCysGlyValle 178
Db 484 GCATGGCGTATTATT 498

RESULT 22
US-09-439-813-1
; Sequence 1, Application US/09439813
; Patent No. 6517845
; GENERAL INFORMATION:
; APPLICANT: Lee, Fang-Jen S.
; APPLICANT: Wu, Chung-Hsiun H.
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS SUPEROXIDE
; TITLE OF INVENTION: DISMUTASE
; FILE REFERENCE: 10457-002001
; CURRENT APPLICATION NUMBER: US/09/439,813
; CURRENT FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 60/108,309
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(720)
US-09-439-813-1

Alignment Scores:
Pred. No.: 2,27e-09 Length: 720
Score: 161.50 Matches: 59
Percent Similarity: 35.81% Conservative: 23
Best Local Similarity: 25.76% Mismatches: 82
Query Match: 17.00% Indels: 65
DB: 4 Gaps: 7

US-10-009-916A-1 (1-180) x US-09-439-813-1 (1-720)
Qy 9 ThrSerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSer 28
Db 43 ACGTCGGTCTGCTCGCGGTGTTTCTGGGCGCGCGTGGCTGCTGAGGCGATGCTCG 102
Qy 29 -----ValThrSerGluVal----- 33
Db 103 TCGCCGCGACGCGCTCTACAGTTCCGGGTACACGCGCGTCTGATTGGACCGGATCGCCC 162
Qy 34 -----HisMetIleAspAsnGlyIleLysGlnSerIleGly 46
Db 163 GCGCGCTGGGACTTTCGGGTCTACGACGAGGATCGCCCGGTGCGCAGAGCTGACCACT 222
Qy 47 ThrValThrPheThrAspThrAspLysGlyLeuGlnIleLys----- 60
Db 223 ACCCTGAGCGGCGCCGACGCGCACGAGTAGCGAGTTCGAGTTCGCCAACCGGC 282
Qy 61 -----ThrAspLeuLysGlyLeuProAlaGlyGluHisGly 72
Db 283 TATGCCACCGTCAAGTCCGCGACGCGCGTCTCGTAAGTCTACGCGCGGTTCACGCGC 342
Qy 73 PheHisIleHisGluGlyGlySerCysGlyPro-----AlaGluHisAspGlyHis 89
Db 343 CTACACATCCACAGGTGGGTAAAGTGTGAGCCCAACTCGGTTGCCCCACCGCGGTGCG 402
```

```
QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLysHisGlu 109
Db 403 CCGCGCAACTTTCTGTCGGCGCGGCCACTACACGTTG-----CCA 444
QY 110 GlyProLeuGlyAsnGlyHisGlyAspLeuProArgLeuValVallysAlaAspGly 129
Db 445 GGGCATACCGGCAACCCCGCAGCGCGACCTGGCTCGGTACGCGGTACGCGGT 504
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThrValVallysGluLeuLysGly 147
Db 505 TCGGCGATGCTGTGACCGACCGAGCGCTTCAACATGACGACCTGTGACGCGCGG 564
QY 148 ---ArgThrValMetIleHisAlaGlyAspAsnTyrSerAspLysPro----- 163
Db 565 AAAACCGCGATCATCTACGCGCGCGCGCACTTTGCCAACATTCGCGCAGACGC 624
QY 164 -----LeuProLeuGlyGly 169
Db 625 TAGCTCCAGGTCAATGGGACTCCGGTCCGAGCAGACGCTTGACCGCGCGCGC 684
QY 170 GlyAlaArgIleAlaCysGlyValIle 178
Db 685 GGCACGCGGTGGCTGGTGTCTATT 711
RESULT 23
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Alignment Scores:
Pred. No.: 0.000772 Length: 4403765
Score: 161.50 Matches: 59
Percent Similarity: 35.81% Conservative: 23
Best Local Similarity: 25.78% Mismatches: 82
Query Match: 17.00% Indels: 65
DB: 3 Gaps: 7
US-10-009-916A-1 (1-180) x US-09-103-840A-2 (1-4403765)
QY 9 ThrSerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSer 28
Db 521082 ACGTGCGTCTCTCCGCGTGTGTTTCTGGGCGCGGTGCGCGTCTGAGCGCATGCTCG 521141
QY 29 -----ValThrSerGluVal----- 33
Db 521142 TCGCGGACGACCGCTCTACAGTTCGGGTACCAACCGCTGATTGGACCGGATCGCCC 521201
QY 34 -----HisMetIleAspAsnGlyIleLysGlnSerIleGly 46
Db 521202 GCGCGTGGGACTTTTCGGGTTCACGACGAGGAGTCCCGCGGTGCGCAGACGCTGACCACT 521261
```

```
QY 47 ThrValThrPheThrAspThrAspLysGlyLeuGlnIleLys----- 60
Db 521262 ACCCTGACGGCGCCCGACGACGAAGGTAGCAGCCGCGAAGTTCGAGTTCCGCAACGCG 521321
QY 61 -----ThrAspLeuLysGlyLeuProAlaGlyGluHisGly 72
Db 521322 TATGCCACCGTCACGATCGGACGACCGGCGTGGTAAGCTACAGCCGCGGTTCACGCG 521381
QY 73 PheHisIleHisGluGlySerCysGlyPro-----AlaGluHisAspGlyHis 89
Db 521382 CTACACATCACCGAGTGGGTAAAGTGTAGCCCACTCGGTTGCCCGCGGTGCG 521441
QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLysHisGlu 109
Db 521442 CCGCGCAACTTCTGTCGGCGCGGCCACTACACGTTG-----CCA 521483
QY 110 GlyProLeuGlyAsnGlyHisGlyAspLeuProArgLeuValVallysAlaAspGly 129
Db 521484 GGGCATACCGGCAACCCCGCAGCGCGACCTGGCTCGGTACGCGGTACGCGGTACGCGT 521543
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThrValVallysGluLeuLysGly 147
Db 521544 TCGGCGATGCTGTGACCGACCGACCGCTTACCATGACGACCTGTGACGCGCGCG 521603
QY 148 ---ArgThrValMetIleHisAlaGlyAspAsnTyrSerAspLysPro----- 163
Db 521604 AAAACCGCGATCATCTACGCGCGCGCGCACTTTGCCAACATTCGCGCAGACGC 521663
QY 164 -----LeuProLeuGlyGly 169
Db 521664 TAGCTCCAGGTCAATGGGACTCCGGTCCGAGCAGACGCTTGACCGCGCGCGCG 521723
QY 170 GlyAlaArgIleAlaCysGlyValIle 178
Db 521724 GGCACGCGGTGGCTGGTGTCTATT 521750
RESULT 24
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 0.000774 Length: 4411529
Score: 161.50 Matches: 59
Percent Similarity: 35.81% Conservative: 23
Best Local Similarity: 25.76% Mismatches: 82
Query Match: 17.00% Indels: 65
DB: 3 Gaps: 7
US-10-009-916A-1 (1-180) x US-09-103-840A-1 (1-4411529)
QY 9 ThrSerIleValThrIleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSer 28
Db 519640 ACGTGCGTCTCTCCGCGTGTGTTTCTGGGCGCGGTGCGCGTCTGAGCGCATGCTCG 519699
```

```
QY 29 -----ValThrSerGluVal----- 33
Db 519700 TCGCCGACACGCGTCTACAGTTCCGGGTACACACCGCGTTCGATTGGACCGCGATCGCCC 519759
QY 34 -----HisMetIleAspAsnGlyIleLysGlnSerIleGly 46
Db 519760 GCGCGGTCCGGACTTTCGGGTCTACGACGAGAGTCCCGGTGCGCAGACGCTGACCACT 519819
QY 47 ThrValThrPheThrAspThrAspLysGlyLeuGlnIleLys----- 60
Db 519820 ACCTGACGCGCGCCGACGCGCAGACGAGGTAGCACCAGCGCGAAGTTCGAGTTCGCCAACGCG 519879
QY 61 -----ThrAspLeuLysGlyLeuProAlaGlyGluHisGly 72
Db 519880 TATGCCACCGTCACGATCGCAGACGCGGTCGCTAAGCTTCACGCGCGCTTCCACGCG 519939
QY 73 PheHisIleHisGluGlySerCysGlyPro-----AlaGluHisAspGlyHis 89
Db 519940 CTACACATCCACGAGTGGGTAAAGTGTGAGCCCAACTCGGTTGCGCCGCCACCGCGGTGCG 519999
QY 90 LeuThrAlaGlyLeuGlnAlaHisGlyHisThrAspProAspLysThrGlyLysHisGlu 109
Db 520000 CCGGCAACTTCTGTCGCGCGCGCCACTACACGTG-----CCA 520041
QY 110 GlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValVallysAlaAspGly 129
Db 520042 GGCATACCGGCACCCCGCCGCGGCGACTTGGCTCGCTGCGAGGTACGCGGTGACGCT 520101
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThrVallysGluIleLysGly----- 147
Db 520102 TCGGCGATGCTGCTGACCAACCGCGCGCTTCCACATGACGACCTGCTGAGCGCGCG 520161
QY 148 ---ArgThrValMetIleHisAlaGlyGlyAspAsnThrSerAspLysPro----- 163
Db 520162 AAAACCGCATCATCATTCACGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 520221
QY 164 -----LeuProLeuGlyGlyGly 169
Db 520222 TAGCTCAGGTCAATGGGACTCCGGGTCCGCGACGAGCAGCTGTGACCCGCGCGCGCC 520281
QY 170 GlyAlaArgIleAlaCysGlyValIle 178
Db 520282 GGCAAGCGGTGGCGTGGCGGTGTCAATT 520308

RESULT 25
US-08-368-236-2
; Sequence 2, Application US/08368236
; Patent No. 5804408
; GENERAL INFORMATION:
; APPLICANT: Hagiwara, et al.
; TITLE OF INVENTION: A METHOD FOR EXPRESSING POLYPEPTIDES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sherman and Shalloway
; STREET: 413 N. Washington Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 Mb STORAGE
; COMPUTER: IBM
; OPERATING SYSTEM: DOS 5.1
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/368,236
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/941,139
; FILING DATE: No. 5804408ember 13, 1992
; APPLICATION NUMBER:
; FILING DATE:
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard Steinberg.
; REGISTRATION NUMBER: 26,588
; REFERENCE/DOCKET NUMBER: S-2336
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 549-2282
; TELEFAX: (703) 836-0106
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1186
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: No
; US-08-368-236-2

Alignment Scores: 7.01e-09 Length: 1186
Pred. No.: 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Score: 160.00 Matches: 52
Percent Similarity: 46.63% Conservative: 24
Best Local Similarity: 31.90% Mismatches: 57
Query Match: 16.84% Indels: 30
DB: 9 Gaps: 9

US-10-009-916A-1 (1-180) x US-08-368-236-2 (1-1186)
QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 427 ACCAAAGCTGTTTGGCTTCTGAAAGGTGACGCGCCCGTTTCAAGGGTATC-----ATC 477
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 478 TTCAACACAGAAAGATCTAACCGTCCGGTTAAAGTTTGGGGTTCTATCAAAAGCGCTGACC 537
QY 68 AlaGlyGluHisGlyPheHisIleHisGlyGlySerCysGlyProAlaGluHisAsp 87
Db 538 GAAGGCTCGCATCGATTCCATGTTTCATGAATTTGGTGAC----- 576
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106
Db 577 -----AACACTGCGAGTTGACCTCTGCGAGGCGCTCATTTCAACCG-----CTGTCGCGT 627
QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuVal 125
Db 628 AAACATGTTGGCGCGAAAGACGACGATGTTGGTGCACCTAGGTAAAGTTACCGCT 687
QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrVallysGlu--- 144
Db 688 GACAAAGACGGTGTGCTGACGTTTCTATCGAAGACTCTGTTATCTCTCTCTCTGTTGAC 747
QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 748 CATGTCATCATCGTCTGCTACTCTGTTGTTTCATGAAAGCGGATGACCTGGGTAAAGGT 807
QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175
Db 808 GGTAAACGAGGAATCTACCAA-----ACCGGTAACGCTGTTCTCTCTGTCGTCATGC 858
QY 176 GlyValIle 178
Db 859 GGTGTTATC 867

RESULT 26
US-08-225-757B-1
; Sequence 1, Application US/08225757B
; Patent No. 5506133
; GENERAL INFORMATION:
; APPLICANT: YU, ET AL.
; TITLE OF INVENTION: Superoxide Dismutase-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
```



```
QY 98 GlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeu---GlyAsnGlyHis 116
Db 493 AATCACTTTAAACCTGAT---GGAGCATCTCATGGGGGCCCCCGAGGACTCTGACCGGCAC 549
QY 117 LysGlyAspLeuProArgLeuValLysAlaAspGlyTleAlaLysGluThrLeuLeu 136
Db 550 CGCGGAGACCTGGCCAAATGTCCGTGTGATGCTGACGGCGCGCCCATCTTTCAGAAATGGAG 609
QY 137 AlaProArgLeuThrValLysGluLleLysGlyArgThrValMetIleHisAlaGlyGly 156
Db 610 GATGACGACGCTGAAGGTGTGGGATGTGATGGCCGAGCCTGATTAATGATGAGGAGAA 669
QY 157 AspAsnTyrSerAspLysProLeuProLeu-----GlyGlyGlyGlyAlaArg 172
Db 670 GATGACCTGGCGGGGAGGCCATCCCTTATCCAAGATCACAGGGAACCTCCGGGGAGAGG 729
QY 173 IleAlaCysGlyValle 178
Db 730 TTGGCCTGTGGCATCAT 747

RESULT 28
US-09-883-985-1
; Sequence 1, Application US/09883985
; Patent No. 6635252
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; ROSEN, CRAIG A.
; FRASER, CLAIRE M.
; GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (cDNA)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..879
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-883-985-1
```

```
Alignment Scores:
Pred. No.: 9,85e-08 Length: 1080
Score: 149.50 Matches: 45
Percent Similarity: 44.52% Conservatives: 20
Best Local Similarity: 30.82% Mismatches: 52
Query Match: 15.74% Indels: 29
DB: 4 Gaps: 6

US-10-009-916A-1 (1-180) x US-09-883-985-1 (1-1080)
QY 46 GlyThrValThrPheThrAsp---ThrAspLysGlyLeuGlnIleLysThrAspLeuLys 64
Db 358 GGGGTGGTGGCTTCCTACAGCTGACCCCTGAGCGCTGCCTCATCGAGGGAACACTATTGAC 417
QY 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGly-----79
Db 418 GGCCCTGGAGCCTGGGCTGTCATGGACTCCACGTCCTACAGTACCGGACCTTACAACAAC 477
QY 80 -----SerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
Db 478 TGCAACAGCTGTGG-----492
QY 98 GlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeu---GlyAsnGlyHis 116
Db 493 AATCACTTTAAACCTGAT---GGAGCATCTCATGGGGGCCCCCGAGGACTCTGACCGGCAC 549
QY 117 LysGlyAspLeuProArgLeuValLysAlaAspGlyTleAlaLysGluThrLeuLeu 136
Db 550 CGCGGAGACCTGGCCAAATGTCCGTGTGATGCTGACGGCGGCCCATCTTTCAGAAATGGAG 609
QY 137 AlaProArgLeuThrValLysGluLleLysGlyArgThrValMetIleHisAlaGlyGly 156
Db 610 GATGACGACGCTGAAGGTGTGGGATGTGATGGCCGAGCCTGATTATGATGAGGAGAA 669
QY 157 AspAsnTyrSerAspLysProLeuProLeu-----GlyGlyGlyGlyAlaArg 172
Db 670 GATGACCTGGCGGGGAGGCCATCCCTTATCCAAGATCACAGGGAACCTCCGGGGAGAGG 729
QY 173 IleAlaCysGlyValle 178
Db 730 TTGGCCTGTGGCATCAT 747

RESULT 29
US-09-248-796A-6301
; Sequence 6301, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6301
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6301

Alignment Scores:
Pred. No.: 5.84e-08 Length: 480
Score: 147.00 Matches: 54
Percent Similarity: 44.05% Conservatives: 20
Best Local Similarity: 32.14% Mismatches: 55
Query Match: 15.47% Indels: 40
DB: 4 Gaps: 8
```





```
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PE/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: PAG1205RP
; US-08-998-416-217

Alignment Scores:
Pred. No.: 0.000274 Length: 747
Score: 117.50 Matches: 31
Percent Similarity: 53.41% Conservative: 16
Best Local Similarity: 35.23% Mismatches: 32
Query Match: 12.37% Indels: 9
DB: 3 Gaps: 4

US-10-009-916A-1 (1-180) x US-08-998-416-217 (1-747)

QY 99 HistyrAspProAspLysThrGlyLysHisGluGlyProLeuGlyLysHisGlyHisLys 117
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 CACTTCACCCATTCAAGAG---AGCAGCGAGTCCGGAGGACGAGAACCGTCAAGTG 62
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 GlyAspLeuProArgLeuValValLysAlaAspGlyLeAlaLysGluThrLeuLeuAla 137
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 GCGGACATGGCAACGTCGCGGACGCAACGGCGTGGCGTAGATCGGCGAAGGAC 122
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 138 ProArgLeuThrVal-----LysGluLeuLysGlyArgThrValMetIleHisAla 154
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 CCTCTAATCAAGATTTTGGTCTCCTAGTCTCGATTCTGGCGCGTACGGTGTGTCCACGCC 182
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 155 GlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGlyGly 170
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 GGCAGAGACACTTAGCGCGCGGCGGCAACGAGGAGTCGCTAAAGCGGCAATGCGGGC 242
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 171 AlaArgIleAlaCysGlyValIle 178
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 CCCAGACCTGCTTGGCGCGTGATT 266
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 32
US-09-248-796A-6300
; Sequence 6300, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6300
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-6300

Alignment Scores:
Pred. No.: 9.02e-05 Length: 291
Score: 116.50 Matches: 30
Percent Similarity: 48.28% Conservative: 12
Best Local Similarity: 34.48% Mismatches: 30
```

```
Query Match: 12.26% Indels: 15
DB: 4 Gaps: 3

US-10-009-916A-1 (1-180) x US-09-248-796A-6300 (1-291)

QY 99 HistyrAspProAspLysThrGlyLysHisGluGlyProLeuGlyLysHisGlyHisLysGly 118
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 40 CATGTGGTTCCAGAGATGATGAAGACATGTTGGTGATTAGGTAAT-----87
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 AspLeuProArgLeuValValLysAlaAspGlyIleAla-----LysGluThrLeu 135
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 88 -----ATTCTACTGATGTTAAGTGTCTAAAGTACCCAAACAGATTTA 135
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 136 LeuAlaProArgLeuThrValLysGluIleLysGlyArgThrValMetIleHisAlaGly 155
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 TTGATTAAATTGATTGGTACAGATTCTATCTTGGGTAGAACTATTGTTGTCATGCTGCT 195
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 156 GlyAspAsnTyrSer-----AspLysProLeuProLeuGlyGlyGlyAla 171
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 196 ACTGATGATTATGGTAAAGTGGTTTTTGAAGATTCTAAAACTACTGTCATGCTGTGCT 255
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 172 ArgIleAlaCysGlyValIle 178
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 256 AGACCTGCTTGTGTGTCATT 276
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 33
5472691-4
; Patent No. 5472691
; APPLICANT: MARKLUND, STEFAN; EDLUND, THOMAS
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,744
; FILING DATE: 24-SP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 897,624
; FILING DATE: 12-JUN-1992
; APPLICATION NUMBER: 576,114
; FILING DATE: 27-aug-1990
; APPLICATION NUMBER: 902,596
; FILING DATE: 02-sep-1986
; SEQ ID NO:4
; LENGTH: 666
5472691-4

Alignment Scores:
Pred. No.: 0.0896 Length: 666
Score: 95.00 Matches: 35
Percent Similarity: 39.23% Conservative: 16
Best Local Similarity: 26.92% Mismatches: 47
Query Match: 10.00% Indels: 32
DB: 6 Gaps: 7

US-10-009-916A-1 (1-180) x 5472691-4 (1-666)

QY 63 LeuLysGlyLeuPro-----AlaGlyGluHisGlyPheHisIleHisGluGly 78
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 CTGAGGGCTTCCGACCGAGCGCAACAGCTCCAGCGCGCCATCCACGTGCACGATTC 300
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 79 GlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly 98
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 GGGGAC-----CTGAGCCAGGGCTCGGAGTCC-----327
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 99 HistyrAspProAspLysThrGlyLysHisGluGlyProLeuGly-----AsnGly 115
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 328 -----ACCGGGCCCCACTACACCCGCTGGCGCTGGCGCCAGCCGAG 369
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAla-----Lys 132
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 370 CACCCGGGCGACTTCGGCACTTCGGGTCCGC---GACGGCAGCCTCTGGAGGTACCGC 426
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 133 GluThrLeuLeuAlaProArgLeuThrValLysGluIleLysGlyArgThrValMetIle 152
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 427 GCCGCCCTGGCCCTCGCTCGCGGCCCGGCACCTCCATCGTGGCGCGCGCTGGTGC 486
QY 153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db 487 CAGCTGGCGAGGACGACCTGGCGCGCGGCACCAACCGCCAGCGCTGGAGAACGGGAAC 546
QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
Db 547 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576

RESULT 34
US-08-556-965-1
; Sequence 1, Application US/08556965
; Patent No. 6025540
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Production of EC-SOD
; NUMBER OF SEQUENCES: 31
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,965
; FILING DATE:
; CLASSIFICATION: 800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..669
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..666
; OTHER INFORMATION: /product= "native EC-SOD"
US-08-556-965-1

Alignment Scores:
Pred. No.: 0.0902 Length: 669
Score: 95.00 Matches: 35
Percent Similarity: 39.23% Conservative: 16
Best Local Similarity: 26.92% Mismatches: 47
Query Match: 10.00% Indels: 32
DB: 3 Gaps: 7

US-10-009-916A-1 (1-180) x US-08-556-965-1 (1-669)
QY 63 LeuLysGlyLeuPro-----AlaGlyGluHisGlyPheHisIleHisGluGly 78
Db 241 CTGGAGGCTTCCCGACCGAGCGGACAGCTCCAGCGCGCATCCACGTGCACCGATTC 300
QY 79 GlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly 98
Db 301 GGGGAC-----ACCGGGCCCACTACAAACCGCTGGCGCTGGCGCTGGCGCTGGCGCT 327
QY 99 HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly-----AsnGly 115
Db 328 -----ACCGGGCCCACTACAAACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCT 369
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAla-----Lys 132
Db 370 CACCGGCGCACTTCGGCAACTTCGCGTCCGC---GACGGCAGCGCTTCGAGGTACCGC 426
QY 133 GluThrLeuLeuAlaProArgLeuThrValLysGluLysGlyArgThrValMetIle 152
```

```
Db 427 GCCGCCCTGGCCCTCGCTCGCGGCCCGGCACCTCCATCGTGGCGCGCGCTGGTGC 486
QY 153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db 487 CAGCTGGCGAGGACGACCTGGCGCGCGGCACCAACCGCCAGCGCTGGAGAACGGGAAC 546
QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
Db 547 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576

RESULT 35
5472691-7
; Patent No. 5472691
; APPLICANT: MARKLUND, STEFAN; EDLUND, THOMAS
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,744
; FILING DATE: 24-SP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 897,624
; FILING DATE: 12-JUN-1992
; APPLICATION NUMBER: 576,114
; FILING DATE: 27-aug-1990
; APPLICATION NUMBER: 902,596
; FILING DATE: 02-sep-1986
; SEQ ID NO: 7:
; LENGTH: 723
5472691-7

Alignment Scores:
Pred. No.: 0.101 Length: 723
Score: 95.00 Matches: 35
Percent Similarity: 39.23% Conservative: 16
Best Local Similarity: 26.92% Mismatches: 47
Query Match: 10.00% Indels: 32
DB: 6 Gaps: 7

US-10-009-916A-1 (1-180) x 5472691-7 (1-723)
QY 63 LeuLysGlyLeuPro-----AlaGlyGluHisGlyPheHisIleHisGluGly 78
Db 295 CTGGAGGCTTCCCGACCGAGCGGACAGCTCCAGCGCGCATCCACGTGCACCGATTC 354
QY 79 GlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly 98
Db 355 GGGGAC-----CTGAGCCAGGGCTGCGAGTCC-----381
QY 99 HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly-----AsnGly 115
Db 382 -----ACCGGGCCCACTACAAACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 423
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAla-----Lys 132
Db 424 CACCGGCGCACTTCGGCAACTTCGCGTCCGC---GACGGCAGCGCTTCGAGGTACCGC 480
QY 133 GluThrLeuLeuAlaProArgLeuThrValLysGluLysGlyArgThrValMetIle 152
Db 481 GCGGCGCTGGCGCTCGCTCGCGCGCGCGCACTCCATCGTGGCGCGCGCTGGTGC 540
QY 153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db 541 CACGCTGGCGAGGACGACCTGGCGCGCGGCGGACCAACCGCCAGCGCTGGAGAACGGGAAC 600
QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
Db 601 GCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 630

RESULT 36
5472691-1
; Patent No. 5472691
; APPLICANT: MARKLUND, STEFAN; EDLUND, THOMAS
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE
```





;  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: U.S.  
; ZIP: 30109-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/908,245  
; FILING DATE: 19920702  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: EMU 111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-815-6508  
; TELEFAX: 404-815-6555  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4089 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Bovine  
; TISSUE TYPE: Aorta  
; CELL TYPE: Endothelial  
; US-07-908-245-1

Alignment Scores:  
Pred. No.: 4.18 Length: 4089  
Score: 90.50 Matches: 31  
Percent Similarity: 43.75% Conservative: 11  
Best Local Similarity: 32.29% Mismatches: 39  
Query Match: 9.53% Indels: 15  
DB: 1 Gaps: 4

US-10-009-916A-1 (1-180) x US-07-908-245-1 (1-4089)

|    |      |  |      |
|----|------|--|------|
| QY | 75   | IleHisGluGlyCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeu       | 94   |
| DB | 1322 | CTTCATGAGGACACCGTGGCGGCGTGGTCCACGATGGTCACCTT---GGCGAGCTG     | 1266 |
| QY | 95   | GlnAlaHisGlyHisThrAspProAspLysThrGlyLysHisGluGlyProLeuGlyAsn | 114  |
| DB | 1265 | AAAGCTGTGCAGCACAGCCAGGTTGATCTCCACGCGCGCTTCTCTCCACAGGGACGA    | 1206 |
| QY | 115  | GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr | 134  |
| DB | 1205 | GGTGGT-----CCGCGTGTCCAGGTCCATGCAGACGCCACATCTCCAGGAT          | 1158 |
| QY | 135  | LeuLeuAlaProArgLeuThrValLysGluIleLysGlyArgThrValMetIleHisAla | 154  |
| DB | 1157 | ATTGTAGCGGTGAGGTGCACA-----CAGGTTCCGCGTGCCCAATCTCCGTGCTCATGTA | 1104 |
| QY | 155  | GlyGlyAspAsnThrSerAspLysProLeuProLeuGlyGlyGly                | 170  |
| DB | 1103 | -----CCAGCCGCTCAAGGGGGCCGCGGA                                | 1080 |

Search completed: November 1, 2004, 20:13:59  
Job time : 1177 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 1, 2004, 06:05:17 ; Search time 1925 Seconds  
(without alignments)

3407.352 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 950

Sequence: 1 MKIKLFFVTSIVTISLLTSL.....DKPLPLGGGARIACGVIPN 180

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US1000916/runat\_26102004\_100223\_1980/app\_query.fasta\_1.327  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1000916 @CGN 1.1 2607 @runat\_26102004\_100223\_1980 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFLOCKS=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 376   | 39.6        | 556    | 8  | AZ302628 GSSBrul56  |
| 2          | 333   | 35.1        | 946    | 8  | AF029454 AF029454   |
| 3          | 307   | 32.3        | 521    | 8  | AZ302772 GSSBrul171 |
| 4          | 213.5 | 22.5        | 587    | 7  | CK432596 UI-D-GC1-  |
| 5          | 212   | 22.3        | 709    | 7  | CK432866 UI-D-GC1-  |
| 6          | 211.5 | 22.3        | 654    | 7  | CF947529 UI-D-GC1-  |
| 7          | 208   | 21.9        | 602    | 7  | CK432358 UI-D-GC1-  |
| 8          | 207.5 | 21.8        | 798    | 3  | CR650645 Tetraodon  |
| 9          | 207.5 | 21.8        | 802    | 6  | CB337089 TC026D04F  |

|    |       |      |      |   |          |           |
|----|-------|------|------|---|----------|-----------|
| 10 | 206.5 | 21.7 | 807  | 3 | CR708484 | Tetraodon |
| 11 | 206.5 | 21.7 | 807  | 3 | CR716180 | Tetraodon |
| 12 | 206.5 | 21.7 | 813  | 3 | CR708907 | Tetraodon |
| 13 | 206.5 | 21.7 | 818  | 3 | CR716473 | Tetraodon |
| 14 | 204.5 | 21.5 | 785  | 5 | BU038660 | DH02G05 H |
| 15 | 204.5 | 21.5 | 775  | 3 | CR710738 | Tetraodon |
| 16 | 204.5 | 21.5 | 780  | 3 | CR706777 | Tetraodon |
| 17 | 204.5 | 21.5 | 781  | 3 | CR713495 | Tetraodon |
| 18 | 204.5 | 21.5 | 782  | 3 | CR712708 | Tetraodon |
| 19 | 204.5 | 21.5 | 784  | 3 | CR707770 | Tetraodon |
| 20 | 204.5 | 21.5 | 784  | 3 | CR710074 | Tetraodon |
| 21 | 204.5 | 21.5 | 790  | 3 | CR708259 | Tetraodon |
| 22 | 204.5 | 21.5 | 791  | 3 | CR705201 | Tetraodon |
| 23 | 204.5 | 21.5 | 793  | 3 | CR711929 | Tetraodon |
| 24 | 204.5 | 21.5 | 794  | 3 | CR708463 | Tetraodon |
| 25 | 204.5 | 21.5 | 794  | 3 | CR712651 | Tetraodon |
| 26 | 204.5 | 21.5 | 797  | 3 | CR710042 | Tetraodon |
| 27 | 204.5 | 21.5 | 798  | 3 | CR705100 | Tetraodon |
| 28 | 204.5 | 21.5 | 798  | 3 | CR709821 | Tetraodon |
| 29 | 204.5 | 21.5 | 798  | 3 | CR712239 | Tetraodon |
| 30 | 204.5 | 21.5 | 804  | 3 | CR714867 | Tetraodon |
| 31 | 204.5 | 21.5 | 820  | 3 | CR710954 | Tetraodon |
| 32 | 203.5 | 21.4 | 761  | 3 | CR719031 | Tetraodon |
| 33 | 203.5 | 21.4 | 763  | 3 | CR715887 | Tetraodon |
| 34 | 203.5 | 21.4 | 763  | 3 | CR719113 | Tetraodon |
| 35 | 203.5 | 21.4 | 765  | 3 | CR715206 | Tetraodon |
| 36 | 203.5 | 21.4 | 775  | 3 | CR705465 | Tetraodon |
| 37 | 203.5 | 21.4 | 780  | 3 | CR715859 | Tetraodon |
| 38 | 203.5 | 21.4 | 781  | 3 | CR717882 | Tetraodon |
| 39 | 203.5 | 21.4 | 783  | 3 | CR713733 | Tetraodon |
| 40 | 203.5 | 21.4 | 790  | 3 | CR705990 | Tetraodon |
| 41 | 202.5 | 21.3 | 786  | 3 | CR714167 | Tetraodon |
| 42 | 202.5 | 21.3 | 1629 | 3 | CR681103 | Tetraodon |
| 43 | 201.5 | 21.2 | 808  | 3 | CR692593 | Tetraodon |
| 44 | 201.5 | 21.2 | 1134 | 3 | CR658900 | Tetraodon |
| 45 | 201.5 | 21.2 | 1359 | 3 | CR660835 | Tetraodon |

#### ALIGNMENTS

RESULT 1  
AZ302628 556 bp DNA linear GSS 06-MAR-2001  
GSSBrul565 Brucella abortus random genomic library Brucella  
melitensis biovar Abortus genomic clone UU1565, genomic survey  
sequence.

LOCUS AZ302628 GI:10128839

DEFINITION GSS.

ACCESSION Brucella melitensis biovar Abortus (Brucella abortus)

VERSION Brucella melitensis biovar Abortus

KEYWORDS Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

SOURCE Brucella.

ORGANISM 1 (bases 1 to 556)

REFERENCE Sanchez,D.O., Zandomeni,R.O., Cravero,S., Verdun,R.E., Pierrou,E.,

Faccio,P., Diaz,G., Lanzavecchia,S., Aguero,F., Frasch,A.C.C.,

Andersson,S.G.E., Rosetti,O.L., Grau,O. and Ugalde,R.A.

Gene discovery through genomic sequencing of Brucella abortus

infect. immun. 69 (2), 865-868 (2001)

JOURNAL 21101034

MEDLINE 11159979

PUBMED

COMMENT

Contact: Siv Andersson

Small Genomes Sequencing Group

Department of Molecular Evolution, Uppsala University

Norbyvägen 18C, S-752 36, Uppsala, Sweden

Tel: 46-18-471-4379

Fax: 46-18-471-6404

Email: Siv.Andersson@ebc.uu.se

Sequences were basecalled with phred and vector was masked with

crossmatch (see http://genome.washington.edu). Sequences were then

trimmed from both ends to remove low quality bases and masked

vector.

Class: shotgun.





Fax: 46-18-471-6404  
 Email: Siv.Andersson@bc.uu.se  
 Sequences were basecalled with phred and vector was masked with  
 crossmatch (see <http://genome.washington.edu>). Sequences were then  
 trimmed from both ends to remove low quality bases and masked  
 vector.

Class: shotgun.

Location/Qualifiers

```

source
1..521
/organism="Brucella melitensis biovar Abortus"
/mol_type="genomic DNA"
/strain="2308"
/db_xref="taxon:235"
/clone="UUI1710"
/clone_lib="Brucella abortus random genomic library"
/note="Vector: modified M13"

```

# ORIGIN

```

Alignment Scores:
Pred. No.: 2,31e-22 Length: 521
Score: 307.00 Matches: 63
Percent Similarity: 65.74% Conservative: 8
Best Local Similarity: 58.33% Mismatches: 30
Query Match: 32.32% Indels: 7
DB: 8 Gaps: 2

```

US-10-009-916A-1 (1-180) x AZ302772 (1-521)

```

QY 80 SerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHis 99
|||||
Db 1 AGCTGCGCTCCGGAGAAAAGCGCAAGATCGTACCGGCTCTTGCTCGCGCGGCAT 60
|||||
QY 100 TyrAspProAspLysThrGlyLysHisGluGlyProLeuGlyAsnGlyHisLysGlyAsp 119
|||||
Db 61 TATGATCGGGTAAATACCATCCATCCATTTAGGACCTGAAGTGATGACATATGGCGCAT 120
|||||
QY 120 LeuProArgLeuValValLysAlaAspGlyLeuAlaLysGluThrLeuLeuAlaProArg 139
|||||
Db 121 TTGCCAGCGCTGAGCGGCAATGCTGAGCGGCAAGGTGAGTGAACCGTTGTGCTCCACAT 180
|||||
QY 140 Leu---ThrValLysGluLeuLysGlyArgThrValMetLeuHisAlaGlyGlyAspAsn 158
|||||
Db 181 CTCAAGAAATTTGGCGGAATCAACACACGCTTTTGTATGTTGCTCCATGTCGAGGGGATAAT 240
|||||
QY 159 TyrSerAspLysProLeuProLeuGlyGlyGlyGlyGlyAlaArgLeuAlaCysGlyValIl 178
|||||
Db 241 TATTCGATTAAGCTTGACCGCTTGTTGGCGGTAGTGAGTGCCCGTTTGCTGCGCGGTAT 300
|||||
QY 178 e-----ProAsn 180
|||||
Db 301 CGAATAACCGCCATGCCCCAAT 322
|||||

```

# RESULT 4

```

CK432596/c 587 bp mRNA linear EST 08-JAN-2004
LOCUS
DEFINITION
UI-D-GC1-aal-p-20-0-UI.s1 UI-D-GC1 Alexandrium tamarense cDNA clone
UI-D-GC1-aal-p-20-0-UI 3', mRNA sequence.

```

CK432596

CK432596.1 GI:40757170

EST.

ORGANISM

Alexandrium tamarense  
 Alexandrium tamarense  
 Eukaryota; Alveolata; Dinophyceae; Gonyaulacales; Gonyaulacaceae;  
 Alexandrium.

1 (bases 1 to 587)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa  
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: Bento-soares@uiowa.edu  
 Tissue Procurement: Provasoli-Guillard National Center for Culture  
 of Marine Phytoplankton (CCMP)  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/dinoflagellate.html>  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

# FEATURES

source

```

Location/Qualifiers
1..587
/organism="Alexandrium tamarense"
/mol_type="mRNA"
/strain="CCMP 1598"
/db_xref="taxon:2926"
/clone="UI-D-GC1-aal-p-20-0-UI"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoR I; Site 2: Not I; UI-D-GC1 is a
normalized library derived from UI-D-GC0. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TACCTCGAGA. Tissue was obtained from the
Provasoli-Guillard National Center for Culture of Marine
Phytoplankton (CCMP).
TAG TISSUE=Alexandrium tamarense
TAG LIB=UI-D-GC1
TAG_SEQ=TACCTCGAGA"

```

# ORIGIN

```

Alignment Scores:
Pred. No.: 2,3e-12 Length: 587
Score: 213.50 Matches: 63
Percent Similarity: 48.45% Conservative: 15
Best Local Similarity: 39.13% Mismatches: 40
Query Match: 22.47% Indels: 43
DB: 7 Gaps: 9

```

US-10-009-916A-1 (1-180) x CK432596 (1-587)

```

QY 46 GlyThrValThrPheThrAspThrAsp---LysGlyLeuGlnLeuLysThrAspLeuLys 64
|||||
Db 561 GGCACCATCACTTTCACGACGCGATGACAGCACTGCACCGCTTGAGTACGAGTCAAG 502
|||||
QY 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAla 84
|||||
Db 501 GGGCTCGCACCGGGGAGCACGCGTCTCCATGATACAGAAAAGGCAGACTTCAGC----- 449
|||||
QY 85 GluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLys 104
|||||
Db 447 -----AATGTTGCGCTAGCGGGGA-----CCCACTACATCCA----- 412
|||||
QY 105 ThrGlyLys---HisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArg 122
|||||
Db 411 TTTGGCAAGTGGCATGCTGCTGCTGATTTGGAACGTCACGTGCGGATCTCGGCAAC 352
|||||
QY 123 LeuValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuVal 142
|||||
Db 351 ATCGTAGCAGACGCGGATGGATTTGCAAGGGCACC-----CTGACCGAC 307
|||||

```

QY 143 LysGluIleLys-----GlyArgThrValMetIleHisAla 154  
 |||||  
 Db 306 AAGTTTCATCAAGATCTTCGGTGAGTACATGTGGTGGCGCTCCATCATGATCATGCA 247  
 |||||  
 QY 155 GlyGlyAspAsnTyr-----GlyArgThrValMetIleHisAla 159  
 |||||  
 Db 246 GACCCTGACGACCTTGGCGGTGGCGACCCCGTGGGCTGGCCCTGAGTGCGCTCCTCTGCC 187  
 |||||  
 QY 160 -----SerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyVal 177  
 |||||  
 Db 186 GCCCAGCCAGCCACACACAAAGACCACCGGCAATGTGGCGCGGCAATTCGATGTGGAGTA 127  
 |||||  
 QY 178 Ile 178  
 |||||  
 Db 126 ATA 124

## RESULT 5

CK432866/c  
 LOCUS  
 DEFINITION UI-D-GC1-aaw-k-15-0-UI.s1 UI-D-GC1 Alexandrium tamarense cDNA clone  
 UI-D-GC1-aaw-k-15-0-UI 3', mRNA sequence.

ACCESSION CK432866  
 VERSION CK432866.1 GI:40757767  
 KEYWORDS EST.  
 SOURCE Alexandrium tamarense  
 ORGANISM Eukaryota; Alveolata; Dinophyceae; Gonyaulacales; Gonyaulacaceae;  
 Alexandrium.

REFERENCE  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548

## COMMENT

Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Provasoli-Guillard National Center for Culture  
 of Marine Phytoplankton (CCMP)  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/dinoflagellate.html>  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

## FEATURES

Location/Qualifiers  
 1..709  
 /organism="Alexandrium tamarense"  
 /mol\_type="mRNA"  
 /strain="CCMP 1598"  
 /db\_xref="taxon:2926"  
 /clone="UI-D-GC1-aaw-k-15-0-UI"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-D-GC1"  
 /note="Vector: pT73-Pac (Pharmacia) with a modified  
 polylinker; Site 1: EcoR I; Site 2: Not I; UI-D-GC1 is a  
 normalized library derived from UI-D-GC0. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 TACCTCGAGA. Tissue was obtained from the

Provasoli-Guillard National Center for Culture of Marine  
 Phytoplankton (CCMP).  
 TAG TISSUE=Alexandrium tamarense  
 TAG LIB=UI-D-GC1  
 TAG\_SEQ=TACCTCGAGA"

## ORIGIN

## Alignment Scores:

Pred. No.: 4.28e-12 Length: 709  
 Score: 212.00 Matches: 69  
 Percent Similarity: 45.03% Conservative: 17  
 Best Local Similarity: 36.13% Mismatches: 57  
 Query Match: 22.32% Indels: 48  
 DB: 7 Gaps: 11

US-10-009-916A-1 (1-180) x CK432866 (1-709)

QY 21 ThrSerValValLeuAlaCysSerValThrSerGluVal-----HisMetIle 36  
 |||||  
 Db 654 ACTGGTGTGTCTATTCCTCCCTCAAGCAACTGTCTGAGTTGGCGAAGCGGGCAAGCGTGC 595  
 |||||  
 QY 37 AspAspAsnGlyIleIysGlnSerIle---GlyThrValThrPheThrAspThrAsp--- 54  
 |||||  
 Db 594 GATGGCTCGATACCAAGACTGGTGTCTACTTGGCACCATCATCTTCCACGACGCGATGCA 535  
 |||||  
 QY 55 LysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHis 74  
 |||||  
 Db 534 GAGAACTGCACCGTTGATTACGAGCTCAAGGGCTCGCACGAGGGGAGGACGCGTTCAT 475  
 |||||  
 QY 75 IleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeu 94  
 |||||  
 Db 474 GTACACGAAAGGCAGACTTCAGC-----AATGGTGGCGCCAGCGCGGA--- 430  
 |||||  
 QY 95 GlnAlaHisGlyHisTyrAspProAspLysThrGlyLys---HisGluGlyProLeuGly 113  
 |||||  
 Db 429 -----CCCCACTACAATCCA-----TTTGGCAAGTGCATGTGGTCTGTATGAT 385  
 |||||  
 QY 114 ---AsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLys 132  
 |||||  
 Db 384 TTGGAACGTGACGTGGCGATCTCGCAACATCGTAGCAGACGCCGCGATGGAATTGGAAG 325  
 |||||  
 QY 133 GluThrLeuAlaProArgLeuThrValLysGluIleLys----- 146  
 |||||  
 Db 324 GGCACC-----ATGCCGACAGTTCATCAAGATCTTCGGTGAGTACACT 280  
 |||||  
 QY 147 -----GlyArgThrValMetIleHisAlaGlyGlyAspAsnTyr----- 159  
 |||||  
 Db 279 GTGGTTGGCCGCTCCATCATGATCATGATGATGATGATGATGATGATGATGATGATGAT 220  
 |||||  
 QY 160 -----SerAspLysProLeuProLeuGly 167  
 |||||  
 Db 219 GTGGGCTGGCGCTGAGGTGGCTCCTCTCGCGGCCCGCCAGCCAGCACAAAGACCGGCG 160  
 |||||  
 QY 168 GlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
 |||||  
 Db 159 AATGCTGGCGCGCGCAATTGCATGTGGAGTAATA 127

## RESULT 6

## CF947529/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## DISCUSSION

## NOTES

## COMMENTS

## BIBLIOGRAPHY

## CROSS-REFERENCES

## OTHER INFORMATION

## REMARKS

## REMARKS

## REMARKS



the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TACCTCGAGA. Tissue was obtained from the Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP).

TAG TISSUE=Alexandrium tanarense

TAG LIB=UI-D-GC1

TAG\_SEQ=TACCTCGAGA"

# ORIGIN

Alignment Scores:  
 Pred. No.: 9,14e-12 Length: 602  
 Score: 208.00 Matches: 60  
 Percent Similarity: 46.82% Conservative: 21  
 Best Local Similarity: 34.68% Mismatches: 44  
 Query Match: 21.89% Indels: 48  
 DB: 7 Gaps: 10

US-10-009-916A-1 (1-180) x CK432358 (1-602)

QY 37 AspAspAsnGlyIleLysGlnSerIle---GlyThrValThrPheThrAspThrAsp--- 54  
 DB 580 GATGCTCAGATPACCAAGACCGGTGCTACTGGCACAATCACATTTCACGACGATGCA 521  
 QY 55 LysGlyLeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHis 74  
 DB 520 GAGTTTATCACCATTGAGTATGACGTCAAGGGGCTCACACAGGGATCACGGCTTCAT 461  
 QY 75 IleHisGluGly------GlySerCysGlyProAlaGluHisAsp 87  
 DB 460 GTGCACGAAAGGCAGATTTCAGCAAGAGGATGCGCGCTGGACCC----- 413  
 QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLys 107  
 DB 412 -----CACTACAATCCA-----TTCGGCAAG 392  
 QY 108 ---HisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125  
 DB 391 TGGCATGGAGCCCTCATGATGAGGCGCTCATCGTTGGGATCTCGGCAACATCACTCGG 332  
 QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrVal----- 142  
 DB 331 GACTCCGATGGAGTTGCGAAGGGGACGATGACGACCAAGTGATCAAGATTTTGGCGAG 272  
 QY 143 LysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyr----- 159  
 DB 271 FACACCGTGTGGCGCGCTCCATCATGATPCCATGACAGCCCGACGACCTTGGCGGTGGC 212  
 QY 160 -----SerAspLysProLeuPro----- 165  
 DB 211 GACCTCGCAGGCTGGCGCTGAGCGGCTCTCTCCCGCCCGCCAGCCGACAGAGACC 152  
 QY 166 LeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
 DB 151 ACGGGCAATGCTGTGTCGGCGCATTCATGTGGAGTGATT 113

# RESULT 8

CR650645 798 bp mRNA linear HTC 11-AUG-2004  
 LOCUS Tetraodon nigroviridis full-length cDNA.  
 DEFINITION

ACCESSION CR650645

VERSION HTC: cDNA; full-length; Tetraodon nigroviridis.

KEYWORDS Tetraodon

SOURCE Tetraodon

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae.

1 (bases 1 to 798)

Genoscope.

REFERENCE

AUTHORS

TITLE

Direct Submission

# JOURNAL

Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage - 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 The sequences are based on single pass reads.  
 More information available at  
 http://www.genoscope.cns.fr/tetraodon.

# COMMENT

Location/Qualifiers

1..798

source

/organism="Tetraodon"

/mol\_type="mRNA"

/db\_xref="taxon:47144"

/tissue\_type="Liver"

# ORIGIN

Alignment Scores:  
 Pred. No.: 1.51e-11 Length: 798  
 Score: 207.50 Matches: 64  
 Percent Similarity: 50.00% Conservative: 24  
 Best Local Similarity: 36.36% Mismatches: 59  
 Query Match: 21.84% Indels: 29  
 DB: 3 Gaps: 10

US-10-009-916A-1 (1-180) x CR650645 (1-798)

QY 22 SerValValLeuAlaCys-----SerValThrSerGluValHisMetIleAsp 37  
 DB 5 AGCGCAGTATTGCCGTGTGTCGTTCTTGAAGTGACACTGCCGAAGATGCTGATAA 64  
 QY 38 Asp-----AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThr 53  
 DB 65 GCTGTTTGGTGTGTTAAAGAGAGCCGGGAGACGACGGAACGGTTTATTTTGAGCAGCAG 124  
 QY 54 AspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHis 71  
 DB 135 GATGAAAGGCTCTGTCAAGTTGACGGGGAGATTAAGGCTCACCGCTGGTAACAC 184  
 QY 72 GlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThr 91  
 DB 185 GGGTTCCATGTCACGCTTTTGGAGAC-----AATACCAATGTTGTCATCAGT 232  
 QY 92 AlaGlyLeuGlnAlaHisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGly 110  
 DB 233 GCAGGC-----CCTCACTACAATCCCAACGACACGAC-----CATGCTGGG 274  
 QY 111 ProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129  
 DB 275 CCTAATGATGAAACAGGCACGCTTGGAGACCTGGGAATGTGACCGCTGGAGCACACGAG 334  
 QY 130 IleAlaLysGluThrLeuAlaProArgLeuThrValLys-----GluIleLys 146  
 DB 335 ATTGCAAGATTGACATAACCGATTTCAGTAATAAACCTCCATGGCAAGTTTCTATAATT 394  
 QY 147 GlyArgThrValMetIleHisAlaGlyLysAspAsnTyr-----SerAspLys 162  
 DB 395 GGCAGAACCATGGTGATCCAGAGAAAGCCCGATGATGCTGGGAAAAAGGAGGCAACGAGAG 454  
 QY 163 ProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
 DB 455 AGCCTTAAACAGGAACGCTGGTGGCGTTGGCTGTGGAGTCATC 502

# RESULT 9

CB337089/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

CB337089.1

GI:40544814

EST.

Tribolium castaneum

Tribolium castaneum (red flour beetle)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Tenebrionidae; Tribolium.

CB337089 802 bp mRNA linear EST 01-JAN-2004

Tc026D04F Tribolium castaneum embryonic cDNA library Tribolium

castaneum cDNA clone Tc026D04 3', mRNA sequence.

```

REFERENCE 1 (bases 1 to 802)
AUTHORS Savard, J. and Tautz, D.
TITLE A Tribolium castaneum EST project
JOURNAL Unpublished (2003)
COMMENT Contact: Savard, J.
Abteilung für Evolutionsgenetik, AG Tautz
Institut für Genetik, Universität zu Köln
Weyertal 121, 50931 Köln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Seq primer: M13P -20.
FEATURES
    source
        1..802
            /organism="Tribolium castaneum"
            /mol_type="mRNA"
            /strain="Wild type"
            /db_xref="taxon:7070"
            /clone="TC026D04"
            /dev_stage="Mixed embryonic stages"
            /note="lib=Tribolium castaneum embryonic cDNA library"
            /vector="pBluescript SK; Site 1: EcoRI; Site 2:
XhoI; Uni-ZAP XR cDNA library (Stratagene) constructed by
Reinhard Schroder (1995)"
ORIGIN
Alignment Scores:
Pred. No.: 1..52e-11 Length: 802
Score: 207.50 Matches: 51
Percent Similarity: 52.45% Conservative: 24
Best Local Similarity: 35.66% Mismatches: 47
Query Match: 21.84% Indels: 21
DB: 6 Gaps: 6

US-10-009-916A-1 (1-180) x CB337089 (1-802)
QY 46 GlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGly 65
Db 602 GCGAAATACCTTACCCAAACCGCAAGAGCGTCCAGTCGAAGCGGTGATCAACGGG 543
QY 66 LeuProAlaGlyGluHisGlyPheHisIleHisGluGlySerCysGlyProAlaGlu 85
Db 542 CTCCTCGAAGGGCAAGCAGCGCTTCCACATCCAGAGAGGGCGCCCTTGGGGACAGTCG 483
QY 86 HisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisThrAspProAspLysThr 105
Db 482 AAGGAC-----GCAGGGGGCCACTTAAACCCGGACAA---450
QY 106 GlyLysHisGluGlyProLeuGlyAsnGly-----HisLysGlyAspLeuProArgLeu 123
Db 449 ---AAAGATCACGGGGCTCTGAGGATGCTGTCGCCACCGTTGTGATCTCGGGAACATC 393
QY 124 ValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLys 143
Db 392 ATCGTGATGATAAAAGTGGGCCATGTTAAATTTCCGACAAAGATCATTTGCTGTAAT 333
QY 144 -----GluIleLysGlyArgThrValMetIleHisAlaGlyGlyAsp-----157
Db 332 GGTGAACACAGTATTATAGGAGGGCTGTGTTGCTCATGAGGGGAGGACGATCTGGG 273
QY 158 -----AsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175
Db 272 AAAGGGAACTTCAATGATGACTCCAAA---ACCACAGGACATGCTGGGGCCAGACTTGTCTGC 216
QY 176 GlyValIle 178
Db 215 GCGCTCATC 207

RESULT 10
CR708484
LOCUS CR708484 807 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR708484

```

```

CR708484.1 GI:51206393
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 807)
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
    source
        1..807
            /organism="Tetraodon"
            /mol_type="mRNA"
            /db_xref="taxon:47144"
            /tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 1..96e-11 Length: 807
Score: 206.50 Matches: 63
Percent Similarity: 51.14% Conservative: 27
Best Local Similarity: 35.80% Mismatches: 57
Query Match: 21.74% Indels: 29
DB: 3 Gaps: 10

US-10-009-916A-1 (1-180) x CR708484 (1-807)
QY 22 SerValValLeuAlaCys-----SerValThrSerGluValHisMetIleAsp 37
Db 5 AGGCCAGTATTGCGGTGTGTGCTTCTTGAAGTGACAACTCGAAGAGTGTGATAAAA 64
QY 38 Asp-----AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThr 53
Db 65 GCTGTTTGGGTGTTAAAGAGGAGCGCGGAGACAGTGGAAACGGTTTATTTTGGACAG 124
QY 54 AspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHis 71
Db 125 GATGAAAGGCTCTCTGCAAGTTGACAGGGAGATTAAAGGGCTGACCGCTGGTGAACAC 184
QY 72 GlyPheHisIleHisGluGlySerCysGlyProAlaGluHisAspGlyHisLeuThr 91
Db 185 GGGTTCATGTCATGCTTTTGGAGAC-----AATACCAATGTTGTCATCAGT 232
QY 92 AlaGlyLeuGlnAlaHisGlyHisThrAspPro---AspLysThrGlyLysHisGluGly 110
Db 233 GCAGGC-----CCTCACTACATCCCAACAAAGACC-----CATGCTGGG 274
QY 111 ProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValLysAlaAspGly 129
Db 275 CCTAACGATGAAACAGGACCGTGTGGAGACCTGGGAAATGTGACCGCTGAAGCAGACCAG 334
QY 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThrValLys-----GluIleLys 146
Db 335 ATGCCCAAGATTGACATAACCGATTCAAGTAATAAGCTCCATGCAAGTTTCTATAATT 394
QY 147 GlyArgThrValMetIleHisAlaGlyGlyAspAsnTyr-----SerAspLys 162
Db 395 GGCAGAACCATGTTGATCCAGAGAGGCTGATGACCTGGGAAAGAGAGCAACGAAGAG 454
QY 163 ProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db 455 AGCCTTAAACAGGAACCGCTGTTGGCGCTTGGCTTGGAGTCAATC 502

RESULT 11
CR716180

```

```

LOCUS      CR716180              807 bp      mRNA      linear      HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION  CR716180
VERSION    CR716180.1  GI:51214419
KEYWORDS   HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE     Tetraodon
            ORGANISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae.
REFERENCE  1 (bases 1 to 807)
AUTHORS   Direct Submission
TITLE     Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL   : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT    (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
            The sequences are based on single pass reads.
            More information available at
            http://www.genoscope.cns.fr/tetraodon.
FEATURES   Location/Qualifiers
            source
            1..807
            /organism="Tetraodon"
            /mol_type="mRNA"
            /db_xref="taxon:47144"
            /tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.:      1.96e-11      Length:      807
Score:          206.50      Matches:      63
Percent Similarity: 51.14%      Conservative: 27
Best Local Similarity: 35.80%      Mismatches: 57
Query Match:    21.74%      Indels:      29
DB:             3          Gaps:      10

US-10-009-916A-1 (1-180) x CR716180 (1-807)

QY      22 SerValValLeuAlaCys-----SerValThrSerGluValHisMetIleAsp 37
Db      5 AGCGCAGTATTCGCGTGTGTCGTTTCTTGAAGTGACAACTGCGAAGATGGTGATAAAA 64
QY      38 Asp-----AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThr 53
Db      65 GCTGTTTGGCTGTTAAAGGACCGCGGGAGACCACTGAGTGAACGGTTATTTTGACGACGAG 124
QY      54 AspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHis 71
Db      125 GATGAAAAGGCTCCTGTCAGAGTTGACAGGGGAGATTAAAGGGCTGACCGCTGGTGAACAC 184
QY      72 GlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThr 91
Db      185 GGGTTCCATGTCCTGCTTTTGGAGAC-----AATACCAATGGTTGCATCAGT 232
QY      92 AlaGlyLeuGlnAlaHisGlyHisTyrrAspPro---AspLysThrGlyLysHisGluGly 110
Db      233 GCAGGC-----CCTCCTACCAATCCCAACCAAGACC-----CATGCTGGG 274
QY      111 ProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129
Db      275 CCTAACGATGAAACAGGCACCTTGGAGACCTGGGAAATGTGACCGCTGAAGCAGACCAAG 334
QY      130 IleAlaLysGluThrLeuAlaProArgLeuThrValLys-----GluIleLys 146
Db      335 ATCGCAAGATTGACATACCAATTCAGTAATAAGCCTCCATGGCAAGTTTCTATAATT 394
QY      147 GlyArgThrValMetIleHisAlaGlyGlyAspAsnTyr-----SerAspLys 162
Db      395 GCGCAACCATGGTGATCCACGAGAGGCTGATGACCTGGGAAAAGGAGGACCAACGAGAG 454
QY      163 ProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db      455 AGCCCTTAAACACAGGAAACGCTGCTGGCGGCTTTGGCGCTGTGGAGTTCATC 502

```

```

RESULT 12
LOCUS    CR708907              813 bp      mRNA      linear      HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION  CR708907
VERSION    CR708907.1  GI:51206816
KEYWORDS   HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE     Tetraodon
            ORGANISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae.
REFERENCE  1 (bases 1 to 813)
AUTHORS   Direct Submission
TITLE     Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL   : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT    (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
            The sequences are based on single pass reads.
            More information available at
            http://www.genoscope.cns.fr/tetraodon.
FEATURES   Location/Qualifiers
            source
            1..813
            /organism="Tetraodon"
            /mol_type="mRNA"
            /db_xref="taxon:47144"
            /tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.:      1.98e-11      Length:      813
Score:          206.50      Matches:      63
Percent Similarity: 51.14%      Conservative: 27
Best Local Similarity: 35.80%      Mismatches: 57
Query Match:    21.74%      Indels:      29
DB:             3          Gaps:      10

US-10-009-916A-1 (1-180) x CR708907 (1-813)

QY      22 SerValValLeuAlaCys-----SerValThrSerGluValHisMetIleAsp 37
Db      12 AGCGCAGTATTCGCGTGTGTCGTTTCTTGAAGTGACAACTGCGAAGATGGTGATAAAA 71
QY      38 Asp-----AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThr 53
Db      72 GCTGTTTGGCTGTTAAAGGACCGCGGGAGACCACTGAGTGAACGGTTATTTTGACGACGAG 131
QY      54 AspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHis 71
Db      132 GATGAAAAGGCTCCTGTCAGAGTTGACGGGGAGATTAAAGGGCTGACCGCTGGTGAACAC 191
QY      72 GlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThr 91
Db      192 GGGTTCCATGTCACCGCTTTTGGAGAC-----AATACCAATGGTTGCATCAGT 239
QY      92 AlaGlyLeuGlnAlaHisGlyHisTyrrAspPro---AspLysThrGlyLysHisGluGly 110
Db      240 GCAGGC-----CCTCCTACCAATCCCAACCAAGACC-----CATGCTGGG 281
QY      111 ProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly 129
Db      282 CCTAACGATGAAACAGGCACCTTGGAGACCTGGGAAATGTGACCGCTGAAGCAGACCAAG 341
QY      130 IleAlaLysGluThrLeuAlaProArgLeuThrValLys-----GluIleLys 146
Db      342 ATCGCAAGATTGACATACCAATTCAGTAATAAGCCTCCATGGCAAGTTTCTATAATT 401
QY      147 GlyArgThrValMetIleHisAlaGlyGlyAspAsnTyr-----SerAspLys 162
Db      402 GCGAAGACCATGGTGATCCACGAGAGCGCCGATGACCTGGGAAAAGGAGGACCAACGAGAG 461

```





```
US-10-009-916A-1 (1-180) x BU038660 (1-765)

QY 1 MetLysIleLeuPhePheValThrSerIleValThrIleSerLeuLeuThrSerIle 20
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 23 ATGTTGTCGCTGGCGATATATATATCTATTTTGTGTAAAGCTACGTAAATTGTGAA 82
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 21 ThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsnGly 40
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 83 ACAGGACTGCTATAGCTGCCTTGTGTCACAGATGTAAC----- 124
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 41 IleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLys 60
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 125 -----GGGAGTATCGAGTTCACGGACAACTGCTGGGCTTCGAGTTACT 169
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 61 ThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySer 80
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 170 GGGAGTATCATTTGGCTGCGCGCGGTAACTACGGGTTCCATGTGCACGAGTTGGGTGAT 229
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 81 CysGlyProAlaGluHisAspGly--HisLeuThrAlaGlyLeuGlnAlaHisGlyHis 99
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 230 ACACCACTTGGATGCATCAGATCGCACTTCAACCTGACGGCAACACTCACGGA--- 286
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 100 TyrAspProAspLysThrGlyLysHisGluGlyProLeuGlyAsnGlyHisLysGlyAsp 119
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 287 ---GGCAGAGATCACACAGTGCAGATGTCGGGACCTCGGAAT----- 328
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 120 LeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArg 139
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 329 ---GTCCTGTCGTAGGAACCTGCGGTTGGAGTGGCTAATGTAGATTTTGTGGACGATGTT 385
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 140 LeuThrValLys-----GluIleLysGlyArgThrValMetIleHisAlaGlyGly 156
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 386 ATTGCACCTCGAGGGCGCAACAGTATTTTGGGCGGTACCTTGGTCTCCACGACGAGAA 445
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 157 AspAsnTyr-----SerAspLysProLeuProLeuGlyGlyGlyAlaArg 172
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 446 GATGATCTAGGCTGGGTAAACAGTGATATTCCTGACTACTGGAACCGGGGCTCGA 505
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 173 IleAlaCysGlyValIle 178
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 506 GTAGCGTGGTGTATT 523
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 15
CR710738
LOCUS
DEFINITION
Tetraodon nigroviridis full-length cDNA.
ACCESSION
CR710738.1 GI:51208647
VERSION
HTC; cDNA; full-length; Tetraodon nigroviridis.
KEYWORDS
Tetraodon
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE
1 (bases 1 to 775)
Genoscope.
Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
COMMENT
Location/Qualifiers
1. .775
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
FEATURES
source
ORIGIN
```

## Alignment Scores:

```
Pred. No.: 3.02e-11 Length: 775
Score: 204.50 Matches: 61
Percent Similarity: 51.46% Conservative: 27
Best local Similarity: 35.67% Mismatches: 60
Query Match: 21.53% Indels: 23
DB: 3 Gaps: 9

US-10-009-916A-1 (1-180) x CR710738 (1-775)

QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 8 TCGTTTCTTGAGTGCACAACTGCGAAGATGCTGATAAAAGCTGTTTGCGTGTTA----- 61
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 62 AAAGAGCGCGGGAGACCGATGGAACGGTTTATTTTGGACGACGATGAAAGGCTCCT 121
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 122 GTCAAGTTGCACAGGGGAGATTAAAGGGCTGACCGCTGGTGAACACGGGTTCCATGCCAT 181
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 182 GCTTTTGGAGAC-----AATACCAATGGTTCATCAGTGCAGGC----- 220
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 221 ---CCTCACTACATCCCCACAAAGACC-----CATGCTGGGCTTAACATGAAAC 271
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 272 AGGCACGTTGGAGACCTGGAAATGTGACCGCTGAAGCAGACAGATCGCCAAGATTGAC 331
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 332 ATAACCGGATTCAGTAATAAGCTCCATGCGCAAGTTCATTAATTTGGCAGAACCATGGTG 391
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 152 IleHisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 392 ATCCACGAGAGGCTGATGACCTGGGAAAGAGGACGACGAGAGAGCCCTTAACACAGGA 451
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 168 GlyGlyAlaArgIleAlaCysGlyValIle 178
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 452 AACGCTGGTGGCGCTTGGCTGTGGAGTCATC 484
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 16
CR706777
LOCUS
DEFINITION
Tetraodon nigroviridis full-length cDNA.
ACCESSION
CR706777
VERSION
HTC; cDNA; full-length; Tetraodon nigroviridis.
KEYWORDS
Tetraodon
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE
1 (bases 1 to 780)
Genoscope.
Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
COMMENT
Location/Qualifiers
1. .780
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
FEATURES
source
```





http://www.genoscope.cns.fr/tetraodon.

# FEATURES

Location/Qualifiers  
1..782  
/organism="Tetraodon"  
/mol\_type="mRNA"  
/db\_xref="taxon:47144"  
/tissue\_type="Eggs"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,06e-11 Length: 782  
Score: 204.50 Matches: 61  
Percent Similarity: 51.46% Conservative: 27  
Best Local Similarity: 35.67% Mismatches: 60  
Query Match: 21.53% Indels: 23  
DB: 3 Gaps: 9

US-10-009-916A-1 (1-180) x CR712708 (1-782)

```

QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38
DB 15 TCGTTTCTTGAGTGCACAACTCGAAGATGGTGATAAAAGCTGTTTGGCGTTA----- 68
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58
DB 69 AAAGAGCGCGGGAGACCACTGCAACGGTATTATTTTGAGCAGCAGGATGAAAAGGCTCCT 128
QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
DB 129 GTCAAGTTGACAGGGGAGATTAAAGCGCTGACCGTGGTGAACACGGGTTCCTATGCCAT 188
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
DB 189 GCTTTTGGAGAC-----AATACCAATGGTTGCATCAGTGCAGGC----- 227
QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
DB 228 ---CCTCACTACAAATCCCAACACAGACC-----CATGCTGGGCCTTAACGATGAAAC 278
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
DB 279 AGGCACGTTGGAGACTGGGAATGTGACCGCTGAAGCAGACACAGATGCCAAGATTGAC 338
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151
DB 339 ATACCGGATTCAATAAGCTCCATGCGCAAGTTCCTATAATTGGCAGAACCATGGTG 398
QY 152 IleHisAlaGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167
DB 399 ATCCACGAGAGGCTGATGACCTGGGAAAGAGGACCAAGAGAGCGCTTAAACACAGA 458
QY 168 GlyGlyGlyAlaArgIleAlaCysGlyValIle 178
DB 459 AACGCTGGTGGCGTTGGCTGTGGAGTCATC 491

```

## RESULT 19

CR707770  
LOCUS CR707770 784 bp mRNA linear HTC 12-AUG-2004  
DEFINITION Tetraodon nigroviridis full-length cDNA.  
ACCESSION CR707770  
VERSION CR707770.1 GI:51205679  
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.  
SOURCE Tetraodon  
ORGANISM Tetraodon

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae.  
1 (bases 1 to 784)

## REFERENCE

1 (bases 1 to 784)  
Direct Submission  
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -  
AUTHORS Genoscope.  
TITLE  
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
More information available at  
http://www.genoscope.cns.fr/tetraodon.

## FEATURES

source  
1..784  
/organism="Tetraodon"  
/mol\_type="mRNA"  
/db\_xref="taxon:47144"  
/tissue\_type="Eggs"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,07e-11 Length: 784  
Score: 204.50 Matches: 61  
Percent Similarity: 51.46% Conservative: 27  
Best Local Similarity: 35.67% Mismatches: 60  
Query Match: 21.53% Indels: 23  
DB: 3 Gaps: 9

US-10-009-916A-1 (1-180) x CR707770 (1-784)

```

QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38
DB 15 TCGTTTCTTGAGTGCACAACTCGAAGATGGTGATAAAAGCTGTTTGGCGTTA----- 68
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58
DB 69 AAAGAGCGCGGGAGACCACTGGAACGGTATTATTTTGAGCAGCAGGATGAAAAGGCTCCT 128
QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
DB 129 GTCAAGTTGACAGGGGAGATTAAAGCGCTGACCGTGGTGAACACGGGTTCCTATGCCAT 188
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
DB 189 GCTTTTGGAGAC-----AATACCAATGGTTGCATCAGTGCAGGC----- 227
QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
DB 228 ---CCTCACTACAAATCCCAACACAGACC-----CATGCTGGGCCTTAACGATGAAAC 278
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
DB 279 AGGCACGTTGGAGACTGGGAATGTGACCGCTGAAGCAGACACAGATGCCAAGATTGAC 338
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151
DB 339 ATACCGGATTCAATAAGCTCCATGCGCAAGTTCCTATAATTGGCAGAACCATGGTG 398
QY 152 IleHisAlaGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167
DB 399 ATCCACGAGAGGCTGATGACCTGGGAAAGAGGACCAAGAGAGCGCTTAAACACAGA 458
QY 168 GlyGlyGlyAlaArgIleAlaCysGlyValIle 178
DB 459 AACGCTGGTGGCGTTGGCTGTGGAGTCATC 491

```

## RESULT 20

CR710074  
LOCUS CR710074 784 bp mRNA linear HTC 12-AUG-2004  
DEFINITION Tetraodon nigroviridis full-length cDNA.  
ACCESSION CR710074  
VERSION CR710074.1 GI:51207983  
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.  
SOURCE Tetraodon  
ORGANISM Tetraodon

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae.  
1 (bases 1 to 784)

## REFERENCE

1 (bases 1 to 784)  
Genoscope.



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Direct Submission  
Submitted (10-AUG-2004) Genoscope - Centre National de Sequençage -  
( 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
More information available at  
http://www.genoscope.cns.fr/tetraodon.

FEATURES  
source

1..791  
Location/Qualifiers  
/organism="Tetraodon"  
/mol\_type="mRNA"  
/db\_xref="taxon:47144"  
/tissue\_type="Eggs"

ORIGIN

Alignment Scores:  
Pred. No.: 3,11e-11 Length: 791  
Score: 204.50 Matches: 61  
Percent Similarity: 51.46% Conservative: 27  
Best Local Similarity: 35.67% Mismatches: 60  
Query Match: 21.53% Indels: 23  
DB: 3 Gaps: 9

US-10-009-916A-1 (1-180) x CR705201 (1-791)

QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38  
Db 6 TCGTTCTTGAAGTCACAACTGCGAAGATGTTGATAAAGCTGTTTGGTGTTA----- 59  
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58  
Db 60 AAAGGAGCGGGGAGACGAGCGGACGGTATTATTGTGAGCAGCATGAAAAGGCTCCT 119  
QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76  
Db 120 GTCAAGTTGACGGGGGAGATTAAAGGCTGACCGCTGGTGAACACGGGTTCCATGTCAC 179  
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96  
Db 180 GCTTTTGAGAC-----AATACCAATGTTGCATCAGTGCAGGC----- 218  
QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114  
Db 219 ---CCTCACTACATCCCAACAGACC-----CATGTGGGCTTAACGATGAAAAC 269  
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134  
Db 270 AGGCACGTGGAGACCTGGGAATGTGACCGCTGAAGCAGCAGCATGCCAAGATTGAC 329  
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151  
Db 330 ATAACCGATTTCAGTAATAAGCTCCATGGCAAGTTTTCTATAATTGGCAGAACCATGGTG 389  
QY 152 IleHisAlaGlyLysAspAsnTyr-----SerAspLysProLeuProLeuGly 167  
Db 390 ATCCACGAGAGGCGGATGACCTGGGAAAAGGAGGACGAAGAGAGCGCTTAAAAACAGGA 449  
QY 168 GlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
Db 450 AACCTGGTGGCGGCTTGGCTGTGGAGTCATC 482

RESULT 23  
CR711929  
LOCUS  
DEFINITION  
CR711929  
ACCESSION  
VERSION  
CR711929.1 GI:51210146

CR711929 793 bp mRNA linear HTC 12-AUG-2004

Tetraodon nigroviridis full-length cDNA.

KEYWORDS  
SOURCE  
ORGANISM

HTC; cDNA; full-length; Tetraodon nigroviridis.  
Tetraodon  
Tetraodon

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae.

REFERENCE  
AUTHORS

1 (bases 1 to 793)

Genoscope.

Direct Submission

Submitted (10-AUG-2004) Genoscope - Centre National de Sequençage -

( 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

More information available at

http://www.genoscope.cns.fr/tetraodon.

FEATURES

source

1..793  
Location/Qualifiers  
/organism="Tetraodon"  
/mol\_type="mRNA"  
/db\_xref="taxon:47144"  
/tissue\_type="Eggs"

ORIGIN

Alignment Scores:  
Pred. No.: 3,12e-11 Length: 793  
Score: 204.50 Matches: 61  
Percent Similarity: 51.46% Conservative: 27  
Best Local Similarity: 35.67% Mismatches: 60  
Query Match: 21.53% Indels: 23  
DB: 3 Gaps: 9

US-10-009-916A-1 (1-180) x CR711929 (1-793)

QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38  
Db 7 TCGTTCTTGAAGTCACAACTGCGAAGATGTTGATAAAGCTGTTTGGTGTTA----- 60  
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58  
Db 61 AAAGGAGCGGGGAGACGAGCGGACGGTATTATTGTGAGCAGCATGAAAAGGCTCCT 120  
QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76  
Db 121 GTCAAGTTGACGGGGGAGATTAAAGGCTGACCGCTGGTGAACACGGGTTCCATGTCAC 180  
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96  
Db 181 GCTTTTGAGAC-----AATACCAATGTTGCATCAGTGCAGGC----- 219  
QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114  
Db 220 ---CCTCACTACATCCCAACAGACC-----CATGTGGGCTTAACGATGAAAAC 270  
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134  
Db 271 AGGCACGTGGAGACCTGGGAATGTGACCGCTGAAGCAGCAGCATGCCAAGATTGAC 330  
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151  
Db 331 ATAACCGATTTCAGTAATAAGCTCCATGGCAAGTTTTCTATAATTGGCAGAACCATGGTG 390  
QY 152 IleHisAlaGlyLysAspAsnTyr-----SerAspLysProLeuProLeuGly 167  
Db 391 ATCCACGAGAGGCGGATGACCTGGGAAAAGGAGGACGAAGAGAGCGCTTAAAAACAGGA 450  
QY 168 GlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
Db 451 AACCTGGTGGCGGCTTGGCTGTGGAGTCATC 483

RESULT 24  
CR708463  
LOCUS

CR708463 794 bp mRNA linear HTC 12-AUG-2004

DEFINITION Tetraodon nigroviridis full-length cDNA.  
 ACCESSION CR708463  
 VERSION CR708463.1 GI:51206372  
 KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.  
 SOURCE Tetraodon  
 ORGANISM Tetraodon  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae.  
 REFERENCE 1 (bases 1 to 794)  
 AUTHORS Direct Submission  
 TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -  
 JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 COMMENT (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 The sequences are based on single pass reads.  
 More information available at  
 http://www.genoscope.cns.fr/tetraodon.  
 FEATURES  
 source  
 1. .794  
 Location/Qualifiers  
 /organism="Tetraodon"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:47144"  
 /tissue\_type="Eggs"  
 ORIGIN  
 Alignment Scores: Length: 794  
 Pred. No.: 3.12e-11 Matches: 61  
 Score: 204.50 Conservative: 27  
 Percent Similarity: 51.46% Mismatches: 60  
 Best Local Similarity: 35.67% Indels: 23  
 Query Match: 21.53% Gaps: 9  
 DB: 3  
 US-10-009-916A-1 (1-180) x CR708463 (1-794)  
 QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38  
 |||||  
 Db 14 TCGTTTCTTGAAGTGCACAACTGCGAGATGGTATATAAAGCTGTTGCGGTGTTA----- 67  
 QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58  
 |||||  
 Db 68 AAGAGCGCGGGGAGATTAAGGGCTGACCGCTGGTGAACACCGGTTCCATGTCAC 127  
 QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGluHisAspGlyHisLeuHis 76  
 |||||  
 Db 128 GTCAAGTTGCGGGGAGATTAAGGGCTGACCGCTGGTGAACACCGGTTCCATGTCAC 187  
 QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96  
 |||||  
 Db 188 GCTTTTGGAGAC-----AATACCAATGGTTGCATCAGTCAGCGC----- 226  
 QY 97 HisGlyHisThrAspPro---AspLysThrGlyLysHisGlyGlyProLeuGly---Asn 114  
 |||||  
 Db 227 ---CCTCACTACATCCCAACACAGACC-----CATGCTGGGCTACAGATGAAAC 277  
 QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134  
 |||||  
 Db 278 AGGCAGCTTGGAGACCTGGGAAATGTGACCGCTGAAGCAGACCATCGCAAGATTGAC 337  
 QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151  
 |||||  
 Db 338 ATAACCGATTTCAGTAATAGCCCTCCATGTCGAAGTTTCTATATTTGCGAAGCATGGTG 397  
 QY 152 IleHisAlaGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167  
 |||||  
 Db 398 ATCCAGAGAGCCCGATGAGCTGGGAAAGAGGAGGACCAAGAGAGCCCTTAAACAGGA 457  
 QY 168 GlyGlyAlaArgIleAlaCysGlyValIle 178  
 |||||  
 Db 458 AACGCTGGTGGGCTTGGCTGTGGAGTCATC 490

RESULT 25  
 CR712651  
 LOCUS Tetraodon nigroviridis full-length cDNA.  
 DEFINITION CR712651  
 ACCESSION CR712651  
 VERSION CR712651.1 GI:51210868  
 KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.  
 SOURCE Tetraodon  
 ORGANISM Tetraodon  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae.  
 REFERENCE 1 (bases 1 to 794)  
 AUTHORS Direct Submission  
 TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -  
 JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 COMMENT (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 The sequences are based on single pass reads.  
 More information available at  
 http://www.genoscope.cns.fr/tetraodon.  
 FEATURES  
 source  
 1. .794  
 Location/Qualifiers  
 /organism="Tetraodon"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:47144"  
 /tissue\_type="Eggs"  
 ORIGIN  
 Alignment Scores: Length: 794  
 Pred. No.: 3.12e-11 Matches: 61  
 Score: 204.50 Conservative: 27  
 Percent Similarity: 51.46% Mismatches: 60  
 Best Local Similarity: 35.67% Indels: 23  
 Query Match: 21.53% Gaps: 9  
 DB: 3  
 US-10-009-916A-1 (1-180) x CR712651 (1-794)  
 QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38  
 |||||  
 Db 14 TCGTTTCTTGAAGTGCACAACTGCGAGATGGTATATAAAGCTGTTGCGGTGTTA----- 67  
 QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58  
 |||||  
 Db 68 AAGAGCGCGGGGAGATTAAGGGCTGACCGCTGGTGAACACCGGTTCCATGTCAC 127  
 QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGluHisAspGlyHisLeuHis 76  
 |||||  
 Db 128 GTCAAGTTGCGGGGAGATTAAGGGCTGACCGCTGGTGAACACCGGTTCCATGTCAC 187  
 QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96  
 |||||  
 Db 188 GCTTTTGGAGAC-----AATACCAATGGTTGCATCAGTCAGCGC----- 226  
 QY 97 HisGlyHisThrAspPro---AspLysThrGlyLysHisGlyGlyProLeuGly---Asn 114  
 |||||  
 Db 227 ---CCTCACTACATCCCAACACAGACC-----CATGCTGGGCTACAGATGAAAC 277  
 QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134  
 |||||  
 Db 278 AGGCAGCTTGGAGACCTGGGAAATGTGACCGCTGAAGCAGACCATCGCAAGATTGAC 337  
 QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151  
 |||||  
 Db 338 ATAACCGATTTCAGTAATAGCCCTCCATGTCGAAGTTTCTATATTTGCGAAGCATGGTG 397  
 QY 152 IleHisAlaGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167  
 |||||  
 Db 398 ATCCAGAGAGCCCGATGAGCTGGGAAAGAGGAGGACCAAGAGAGCCCTTAAACAGGA 457  
 QY 168 GlyGlyAlaArgIleAlaCysGlyValIle 178  
 |||||

```
Db 458 AACGCTGGTGGCGTTTGGCTGTGGAGTCATC 490
|||||
RESULT 26
LOCUS CR710042 797 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR710042
VERSION CR710042.1 GI:51207951
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 797)
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..797
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3,14e-11 Length: 797
Score: 204.50 Matches: 61
Percent Similarity: 51.46% Conservative: 27
Best Local Similarity: 35.67% Mismatches: 60
Query Match: 21.53% Indels: 23
Gaps: 9
DB: 3
US-10-009-916A-1 (1-180) x CR710042 (1-797)
QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38
Db 17 TCGTTTCTGAAGTGACAACTGCGAAGATGGTGATAAAAGCTGTTTGGTGTTA----- 70
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58
Db 71 AAAGGAGCGGGGAGACCGAGCGGAAACGGTTTATTTTGAGCAGCAGGATGAAAGGCTCCT 130
QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
Db 131 GTCAAGTTGACGGGGGAGATTAAAGGGCTGACCGCTGGTGAACACGGGTTCCATGTCAC 190
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db 191 GCTTTTGAGAC-----AATACCAATGGTTGCATCAGTGCAGGC----- 229
QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
Db 230 ---CCTCACTACAATCCCAACAAAGACC-----CATGCTGGGCCTTAACGATGAAAC 280
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
Db 281 AGGCACGTTGGAGACTGGGAATGTGACCGCTGAAGCAGACCATGCCAGATTGAC 340
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151
Db 341 ATAACCGATTTCAGTAATAGCTCCATGCGCAAGTTTCTATAAATGGCAGAACCATGTTG 400
QY 152 IleHisAlaGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167
|||||
|||||
Db 401 ATCCACGAGAAGCGCGATGACCTGGAAAAAGGAGCAAGAGAGAGCTTAAAAACAGGA 460
168 GlyGlyGlyAlaArgIleAlaCysGlyValIle 178
461 AACGCTGGTGGCGTTTGGCTGTGGAGTCATC 493
|||||
RESULT 27
LOCUS CR705100 798 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR705100
VERSION CR705100.1 GI:51203009
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 798)
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..798
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3,14e-11 Length: 798
Score: 204.50 Matches: 61
Percent Similarity: 51.46% Conservative: 27
Best Local Similarity: 35.67% Mismatches: 60
Query Match: 21.53% Indels: 23
Gaps: 9
DB: 3
US-10-009-916A-1 (1-180) x CR705100 (1-798)
QY 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38
Db 14 TCGTTTCTGAAGTGACAACTGCGAAGATGGTGATAAAAGCTGTTTGGTGTTA----- 67
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58
Db 68 AAAGGAGCGGGGAGACCGAGTGAACGGTTTATTTTGAGCAGCAGGATGAAAGGCTCCT 127
QY 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76
Db 128 GTCAAGTTGACAGGGGAGATTAAAGGGCTGACCGCTGGTGAACACGGGTTCCATGTCAT 187
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db 188 GCTTTTGAGAC-----AATACCAATGGTTGCATCAGTGCAGGC----- 226
QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
Db 227 ---CCTCACTACAATCCCAACAAAGACC-----CATGCTGGGCCTTAACGATGAAAC 277
QY 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
Db 278 AGGCACGTTGGAGACCTGGGAATGTGACCGCTGAAGCAGACCATGCCAGATTGAC 337
QY 135 LeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151
Db 338 ATAACCGATTTCAGTAATAGCTCCATGCGCAAGTTTCTATAAATGGCAGAACCATGTTG 397
|||||
```

|   |   |  |
|---|---|--|
| QY  | 152   | ileHisalaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167          |
| Db  | 398   | ATCCACGAGAAAGCGTGATGACCTGGGAAAAAGGAGGCAACGAGAGACGCTTTAAAAACGGA 457 |
| QY  | 168   | GlyGlyGlyAlaArgIleAlaCysGlyValIle 178                              |
| Db  | 458   | AACGCTGGTGGCGCTTTGGCTGTGGAGTCATC 490                               |
| RESULT 28                                   |   |  |
| LOCUS                                       | CR709821  | 798 bp mRNA linear HTC 12-AUG-2004                                 |
| DEFINITION                                  | Tetraodon nigroviridis full-length cDNA.                            |  |
| ACCESSION                                   | CR709821  |  |
| VERSION                                     | CR709821.1 GI:51207730  |  |
| KEYWORDS                                    | HTC; cDNA; full-length; Tetraodon nigroviridis.                     |  |
| SOURCE                                      | Tetraodon   |  |
| ORGANISM                                    | Tetraodon   |  |
| REFERENCE                                   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |  |
| AUTHORS                                     | Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  |  |
| TITLE                                       | Acanthomorphia; Acanthopterygii; Percomorpha; Tetraodontiformes;    |  |
| JOURNAL                                     | Tetraodontidae; Tetraodontidae.                                     |  |
| COMMENT                                     | 1 (bases 1 to 798)  |  |
| FEATURES                                    | Genoscope.  |  |
| SOURCE                                      | Direct Submission   |  |
|   | Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage - |  |
|   | : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE        |  |
|   | (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)     |  |
|   | The sequences are based on single pass reads.                       |  |
|   | More information available at                                       |  |
|   | http://www.genoscope.cns.fr/tetraodon.                              |  |
|   | Location/Qualifiers   |  |
|   | 1..798  |  |
|   | /organism="Tetraodon"   |  |
|   | /mol_type="mRNA"  |  |
|   | /db_xref="taxon:47144"  |  |
|   | /tissue_type="Eggs"   |  |
| ORIGIN                                      |   |  |
| Alignment Scores:                           |   |  |
| Pred. No.:                                  | 3.14e-11  | Length: 798  |
| Score:                                      | 204.50  | Matches: 61  |
| Percent Similarity:                         | 51.46%  | Conservative: 27   |
| Best Local Similarity:                      | 35.67%  | Mismatches: 60   |
| Query Match:                                | 21.53%  | Indels: 23   |
| DB:   | 3   | Gaps: 9  |
| US-10-009-916A-1 (1-180) x CR709821 (1-798) |   |  |
| QY  | 19  | SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsp 38    |
| Db  | 14  | TCGTTTCTTGAATGACAACTGCGAATGGTGATATAAAGCTGTTTCGGTGTTA----- 67       |
| QY  | 39  | AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58    |
| Db  | 68  | AAAGGACCGGGGAGACACCGGAACGGTTTATTTCGAGCAGCAGGATGAAGAGCTCCT 127      |
| QY  | 59  | IleIysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76     |
| Db  | 128   | GTCAGTTGACGGGGGAGATTAAGGGCTGACCGCTGGTGAACACGGGTTCCATGTCCAC 187     |
| QY  | 77  | GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96    |
| Db  | 188   | GCITTTGGACAC-----AATACCAATGTTTGCAATCAGTCAGTCAGGC----- 226          |
| QY  | 97  | HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114   |
| Db  | 227   | ---CCTCACTACAATCCCCACACAAGACC-----CATGCTGGCCCTAACGATGAAC 277       |
| QY  | 115   | GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134   |
| Db  | 278   | AGGCACCTTGGACCTTGGGAATGTGACCGCTGAAGCAGCACCAGATCGCAAGATTGAC 337     |

|   |   |   |                 |
|---|---|---|-----------------|
| Qy  | 135   | LeuLeuAlaProArgLeuThrValLys-----GlulleLysGlyArgThrValMet        | 151             |
| At  | 135   | LeuLeuAlaProArgLeuThrValLys-----GlulleLysGlyArgThrValMet        | 151             |
| Db  | 338   | ATACCGATTTCAGTAATAAGCCCTCATGGCAAGTATTTTCTATAATTGCGAAGAACCATGGTG | 397             |
| Qy  | 152   | IleHisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGly           | 167             |
| Db  | 398   | ATCCACAGAGAGCCGATGACCTGGGNAAGGAGGACCAAGAGAGCCCTTAAACACAGGA      | 457             |
| Qy  | 168   | GlyGlyAlaArgIleAlaCysGlyValIle                                  | 178             |
| Db  | 458   | AACGCTGGTGGCGTTTGGCCCTGTGGAGTCATC                               | 490             |
| RESULT                                      | 29  |   |                 |
| LOCUS                                       | CR712239  | 798 bp mRNA linear  | HTC 12-AUG-2004 |
| DEFINITION                                  | Tetraodon nigroviridis full-length cDNA.  |   |                 |
| ACCESSION                                   | CR712239  |   |                 |
| VERSION                                     | CR712239.1  | GI:51210456   |                 |
| KEYWORDS                                    | HTC; cDNA; full-length; Tetraodon nigroviridis.   |   |                 |
| SOURCE                                      | Tetraodon   |   |                 |
| ORGANISM                                    | Tetraodon   |   |                 |
| REFERENCE                                   | 1 (bases 1 to 798)  |   |                 |
| AUTHORS                                     | Genoscope.  |   |                 |
| TITLE                                       | Direct Submission   |   |                 |
| JOURNAL                                     | Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage   |   |                 |
| COMMENT                                     | : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE<br>(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)<br>The sequences are based on single pass reads.<br>More information available at<br>http://www.genoscope.cns.fr/tetraodon. |   |                 |
| FEATURES                                    | source  |   |                 |
|   | Location/Qualifiers   |   |                 |
|   | 1..798  |   |                 |
|   | /organism="Tetraodon"   |   |                 |
|   | /mol_type="mRNA"  |   |                 |
|   | /db_xref="taxon:47144"  |   |                 |
|   | /tissue_type="Eggs"   |   |                 |
| ORIGIN                                      |   |   |                 |
| Alignment Scores:                           |   |   |                 |
| Pred. No.:                                  | 3,14e-11  | Length:   | 798             |
| Score:                                      | 204.50  | Matches:  | 61              |
| Percent Similarity:                         | 51.46%  | Conservative:   | 27              |
| Best Local Similarity:                      | 35.67%  | Mismatches:   | 60              |
| Query Match:                                | 21.53%  | Indels:   | 23              |
| DB:   | 3   | Gaps:   | 9               |
| US-10-009-916A-1 (1-180) x CR712239 (1-798) |   |   |                 |
| Qy  | 19  | SerIleThrSerValValLeuAlaCysSerValThrSerCluValHisMetIleAspAsp    | 38              |
| Db  | 14  | TCGTTTCTTGAATGACAACTGCGAAGATGGTATAAAGCTGTTTCGGTGTTA-----        | 67              |
| Qy  | 39  | AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln    | 58              |
| Db  | 68  | AAAGAGACGGGGAGACCACTGGAACGGTTATTTTGGACGACGAGATGAAGAGGTCCT       | 127             |
| Qy  | 59  | IleLysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis     | 76              |
| Db  | 128   | GTCAAGTTGACAGGGGAGATTAAGAGGGCTGACCGCTGGTGAACACGCGTTCCATGTCAT    | 187             |
| Qy  | 77  | GlulGlySerCysGlyProAlaGluHisAspGlyHisLeuThrIleGlyLeuGlnAla      | 96              |
| Db  | 188   | GCCTTTGGAGAC-----AATACCAATGGTTGCATCAGTSCAGGC-----               | 226             |
| Qy  | 97  | HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluCluProLeuGly---Asn    | 114             |
| Db  | 227   | ---CCTCACTACAAATCCCAACAGACC-----CATGCTGGGCTTAACGATGAAC          | 277             |
| Qy  | 115   | GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr    | 134             |







```

QY 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 ---CCTCACTACAAATCCCAACAAGACC-----CATGCTGGGCTTAACGATGAAGAAC 280

QY 115 GlyHisGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 AGCAGCTTGGAGACCTGGGAATGTGACCGCTGAAGCAGACCAAGATCGCAAGATTGAC 340

QY 135 LeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMet 151
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 ATACCGATTCAAGTAATAGCCCTCCATGGCAAGTTTCTATATGTCAGCAACCATGGTG 400

QY 152 IleHisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 ATCCACGAGAAGCGCGATGATCGCTGGGAAAAGGAGGCAACGAAGAGAGCCTTAAACAGGA 460

QY 168 GlyGlyAlaArgIleAlaCysGlyValIle 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 AACGCTGGTGACGCTTGGCCTGTGGAGTCATC 493

RESULT 32
CR719031
LOCUS CR719031 761 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR719031
VERSION CR719031.1 GI:51217282
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 761)
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
[E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..761
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3,77e-11 Length: 761
Score: 203.50 Matches: 56
Percent Similarity: 54.00% Conservative: 25
Best Local Similarity: 37.33% Mismatches: 48
Query Match: 21.42% Indels: 21
DB: 3 Gaps: 8

US-10-009-916A-1 (1-180) x CR719031 (1-761)
QY 40 GlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIle 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49 GGAGCGGGGAGACCACTGGAACCGTTATTTTGGACGACGAGATGAAGGCTCTCTGTC 108

QY 60 LysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 AAGTTGACAGGGGAGATTAAGGGCTGACCGCTGGTGAACACCGGTTCCATGTCATGCT 168

QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 TTTGGAGAC-----AATACCAATGGTTGCATCAGTCAGGC----- 204

```

```

QY 98 GlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 CCTCACTACAAATCCCAACAAGACC-----CATGCTGGGCTTAACGATGAAGAACAGG 258

QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 CAGCTTGGAGACCTGGGAATGTGACCGCTGAAGCAGACCAAGATCGCAAGATTGACATA 318

QY 136 LeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrValMetIle 152
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 ACGGATTCAAGTAATAGCCCTCCATGGCAAGTTTCTATATGTCAGCAACCATGGTGATC 378

QY 153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 CACGAGAAGCGCTGATGACCTGGGAAAAGGAGGCAACGAAGAGAGCCTTAAACAGGAAC 438

QY 169 GlyGlyAlaArgIleAlaCysGlyValIle 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 GCTGGTGGGCGTTTGGCCTGTGGAGTCATC 468

RESULT 33
CR715887
LOCUS CR715887 763 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR715887
VERSION CR715887.1 GI:51214121
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 763)
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
[E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..763
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3,78e-11 Length: 763
Score: 203.50 Matches: 56
Percent Similarity: 54.00% Conservative: 25
Best Local Similarity: 37.33% Mismatches: 48
Query Match: 21.42% Indels: 21
DB: 3 Gaps: 8

US-10-009-916A-1 (1-180) x CR715887 (1-763)
QY 40 GlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIle 59
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 GGAGCGGGGAGACCACTGGAACCGTTATTTTGGACGACGAGATGAAGGCTCTCTGTC 112

QY 60 LysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 AAGTTGACAGGGGAGATTAAGGGCTGACCGCTGGTGAACACCGGTTCCATGTCACGCT 172

QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 TTTGGAGAC-----AATACCAATGGTTGCATCAGTCAGGC----- 208

QY 98 GlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115

```

```
Db 209 CCTCACTACAATCCCAACAAGACC-----CATGCTGGGCGCTTAACGATGAACAGG 262
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyLeuAlaLysGluThrLeu 135
Db 263 CACGTTGGAGACCTGGGAAATGTGACCGCTGAAGCTGACGACGATGCCAAGATTGACATA 322
QY 136 LeuAlaProArgLeuThrValLys-----GluLeLysGlyArgThrValMetile 152
Db 323 ACCGATTTCAGTAATAAGCCTCCATGCGCAAGTTTCTATAATTGGCAGAACCATGTGATC 382
QY 153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db 383 CACGAGAAGCGCGATGACCTGGGAAAGAGGAGCAACGAAGAGAGCCTTAAACACAGGAAC 442
QY 169 GlyGlyAlaArgIleAlaCysGlyValile 178
Db 443 GCTGTGGCGGTTTGGCCTGTGGAGTCATC 472

RESULT 34
CR719113
LOCUS CR719113 763 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR719113
VERSION CR719113.1 GI:51217364
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 763)
AUTHORS Genoscope.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
location/Qualifiers
1..763
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3.78e-11 Length: 763
Score: 203.50 Matches: 56
Percent Similarity: 54.00% Conservative: 25
Best Local Similarity: 37.33% Mismatches: 48
Query Match: 21.42% Indels: 21
DB: 3 Gaps: 8
US-10-009-916A-1 (1-180) x CR719113 (1-763)
QY 40 GlyLeLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIle 59
Db 41 GGAGCGGGGAGACCACTGGAACGGTTATTATTGGACGACGAGATGAAGGCTCCTGTC 100
QY 60 LysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
Db 101 AAGTTGACAGGGGAGATTAAAGGCTGACCGCTGTTGAACACGGGTTCCATGTCATGCT 160
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
Db 161 TTTGGAGAC-----AATACCAATGTTGCATCAGTCAGGC----- 196
QY 98 GlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
Db 115 TTTGGAGAC-----AATACCAATGTTGCATCAGTCAGGC----- 196
```

```
Db 197 CCTCACTACAATCCCAACAAGACC-----CATGCTGGGCGCTTAACGATGAACAGG 250
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyLeuAlaLysGluThrLeu 135
Db 251 CACGTTGGAGACCTGGGAAATGTGACCGCTGAAGCAGCACGATGCCAAGATTGACATA 310
QY 136 LeuAlaProArgLeuThrValLys-----GluLeLysGlyArgThrValMetile 152
Db 311 ACCGATTTCAGTAATAAGCCTCCATGCGCAAGTTTCTATAATTGGCAGAACCATGTGATC 370
QY 153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db 371 CACGAGAAGCGCTGATGACCTGGGAAAGAGGAGCAACGAAGAGAGCCTTAAACACAGGAAC 430
QY 169 GlyGlyAlaArgIleAlaCysGlyValile 178
Db 431 GCTGTGGCGGTTTGGCCTGTGGAGTCATC 460

RESULT 35
CR715206
LOCUS CR715206 765 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR715206
VERSION CR715206.1 GI:51213423
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 765)
AUTHORS Genoscope.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
location/Qualifiers
1..765
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3.79e-11 Length: 765
Score: 203.50 Matches: 56
Percent Similarity: 54.00% Conservative: 25
Best Local Similarity: 37.33% Mismatches: 48
Query Match: 21.42% Indels: 21
DB: 3 Gaps: 8
US-10-009-916A-1 (1-180) x CR715206 (1-765)
QY 40 GlyLeLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIle 59
Db 41 GGAGCGGGGAGACCACTGGAACGGTTATTATTGGACGACGAGATGAAGGCTCCTGTC 100
QY 60 LysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
Db 101 AAGTTGACAGGGGAGATTAAAGGCTGACCGCTGTTGAACACGGGTTCCATGTCATGCT 160
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
Db 161 TTTGGAGAC-----AATACCAATGTTGCATCAGTCAGGC----- 196
QY 98 GlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
Db 197 CCTCACTACAATCCCAACAAGACC-----CATGCTGGGCGCTTAACGATGAACAGG 250
```



```

Db      256 CACGTGTGGAGACCTGGGAAATGTGACCGCTGAAGCAGACCATCGCCAAAGATTGACATA 315
QY      136 LeuAlaProArgLeuThrVallys-----GlulleLysGlyArgThrValMetile 152
Db      316 ACCGATTTCAGTAATAAGCGCTCATGGCAAGTTTCTATAAATGGCAGAACCATGTGTGATC 375
QY      153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db      376 CACGAGAAGGCGGTGTTGGCCTGTGGAGTCATC 465
QY      169 GlyGlyAlaArgIleAlaCysGlyValile 178
Db      436 GCTGTGTGGCGTGTGGCCTGTGGAGTCATC 465

RESULT 39
LOCUS   CR717882 781 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR717882
VERSION CR717882.1 GI:51216133
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 781)
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
location/Qualifiers
1..781
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3,9e-11 Length: 781
Score: 203.50 Matches: 56
Percent Similarity: 54.00% Conservative: 25
Best Local Similarity: 37.33% Mismatches: 48
Query Match: 21.42% Indels: 21
DB: 3 Gaps: 8
US-10-009-916A-1 (1-180) x CR717882 (1-781)
QY 40 GlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIle 59
Db 44 GGAGCGGGGAGACGACGCGAAGCGTTTATTTTGGACGACGAGATGAAAAGGCTCCTGTC 103
QY 60 LysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
Db 104 AAGTTGACGGGGGAGATTAAAGGGCTGACCGCTGTGTGAACACGGGTTCCATGTCACGCT 163
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
Db 164 TTTGGAGAC-----AATACCAATGGTTGCATCAGTCAGGC----- 199
QY 98 GlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
Db 200 CCTCACTACAATCCCAACAAGACG-----CATGCTGGGCGCTAAACGATGAAAACAGG 253
QY 116 HisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGluThrLeu 135

```

```

Db      254 CACGTGTGGAGACCTGGGAAATGTGACCGCTGAAGCTGACCATCGCCAAAGATTGACATA 313
QY      136 LeuAlaProArgLeuThrVallys-----GlulleLysGlyArgThrValMetile 152
Db      314 ACCGATTTCAGTAATAAGCGCTCATGGCAAGTTTCTATAAATGGCAGAACCATGTGTGATC 373
QY      153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db      374 CACGAGAAGGCGGTGACCTGGGAAAGGAGGCAAGAGAGAGCTTTAAAACACGGAAC 433
QY      169 GlyGlyAlaArgIleAlaCysGlyValile 178
Db      434 GCTGTGTGGCGTGTGGCCTGTGGAGTCATC 463

RESULT 39
LOCUS   CR713733 783 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR713733
VERSION CR713733.1 GI:51211950
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 783)
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
location/Qualifiers
1..783
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"
ORIGIN
Alignment Scores:
Pred. No.: 3,91e-11 Length: 783
Score: 203.50 Matches: 56
Percent Similarity: 54.00% Conservative: 25
Best Local Similarity: 37.33% Mismatches: 48
Query Match: 21.42% Indels: 21
DB: 3 Gaps: 8
US-10-009-916A-1 (1-180) x CR713733 (1-783)
QY 40 GlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIle 59
Db 49 GGAGCGGGGAGACGACGCGAAGCGTTTATTTTGGACGACGAGATGAAAAGGCTCCTGTC 108
QY 60 LysThr-----AspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
Db 109 AAGTTGACGGGGGAGATTAAAGGGCTGACCGCTGTGTGAACACGGGTTCCATGTCACGCT 168
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
Db 169 TTTGGAGAC-----AATACCAATGGTTGCATCAGTCAGGC----- 204
QY 98 GlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
Db 205 CCTCACTACAATCCCAACAAGACG-----CATGCTGGGCGCTAAACGATGAAAACAGG 258
QY 116 HisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGluThrLeu 135
Db 259 CACGTGTGGAGACCTGGGAAATGTGACCGCTGAAGCTGACCATCGCCAAAGATTGACATA 318

```

```

Qy 136 LeuAlaProArgLeuThrValLys-----GluLeuLysGlyArgThrValMetile 152
Db 319 ACCGATTTCAGTAATAAGCCTCCATGGCAAGTTTCTATAATTGGCAGAACCATGGTGCATC 378
Qy 153 HisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGlyGly 168
Db 379 CACGAGAGCCGATGACCTGGGAAAAGGAGGACCAAGAGAGAGCCTTAAACAGGAAAC 438
Qy 169 GlyGlyAlaArgIleAlaCysGlyValile 178
Db 439 GCTGGTGGCGCTTGGCTGTGGAGTCATC 468

RESULT 40
CR705990
LOCUS
DEFINITION
Tetraodon nigroviridis full-length cDNA.
ACCESSION
CR705990
VERSION
CR705990.1 GI:51203899
KEYWORDS
HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE
Tetraodon
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
1 (bases 1 to 790)
Genoscope.
Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
location/Qualifiers
1..790
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eggs"

ORIGIN
Alignment Scores:
Pred. No.: 3..96e-11 Length: 790
Score: 203.50 Matches: 61
Percent Similarity: 51.46% Conservative: 27
Best Local Similarity: 35.67% Mismatches: 60
Query Match: 21.42% Indels: 23
DB: 3 Gaps: 9

US-10-009-916a-1 (1-180) x CR705990 (1-790)
Qy 19 SerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMetileAspAsp 38
Db 6 TCGTTTCTTGAAGTGACAACTGCCAGATGGTGATATAAAGCTGTTTCGCTGTTA----- 59
Qy 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGln 58
Db 60 AAAGAGCCGGGAGACACCGAAGCGTTATTTTGAGCAGCAGGATGAAGAGGCTCCT 119
Qy 59 IleLysThr-----AspLeuLysGlyLeuProAlaGlyLysGlyPheHisIleHis 76
Db 120 GTCAGTTGACGGGGGAGATTAAAGGCTGACCGCTGGTGAAACACCGGTTCCATGTCAC 179
Qy 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAla 96
Db 180 GCTTTTGGAGAC-----AATACCAATGGTTGATCATGTCAGCAGC----- 218
Qy 97 HisGlyHisTyrAspPro---AspLysThrGlyLysHisGluGlyProLeuGly---Asn 114
Db 219 ---CCTCCTACATAATCCCAACAAAGACC-----CATGTTGGGCGCTAACGATGAAGAAC 269

```

```

Qy 115 GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThr 134
Db 270 AGGCACGTTGGAGACCTGGGAAATGTGACCGCTGAAGCAGACCCAGATCGCAAGATTGAC 329
Qy 135 LeuLeuAlaProArgLeuThrValLys-----GluLeuLysGlyArgThrValMet 151
Db 330 ATAACCGATTTCAGTAATAAGCCTCCATGGCAAGTTTCTATAATTGGCAGAACCATGGTG 389
Qy 152 IleHisAlaGlyGlyAspAsnTyr-----SerAspLysProLeuProLeuGly 167
Db 390 ATCCACGAGAGCCGATGACCTGGGAAAAGGAGGCAACGAGAGAGCCTTAAACAGGA 449
Qy 168 GlyGlyAlaArgIleAlaCysGlyValile 178
Db 450 AACGCTGGTGGCGCTTGGCTGTGGAGTCATC 482

```

Search completed: November 2, 2004, 02:24:39  
Job time : 1934 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 31, 2004, 05:33:32 ; Search time 1664 Seconds  
(without alignments)  
554.668 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 950

Sequence: 1 MKIKLFFVTSIVTISLLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US1009916/runat\_26102004\_100225\_2092/app\_query.fasta\_1.327  
-DB=Published Applications NA -QFWT=fastap -SUFFIX=rnpb -MINWATCH=0.1  
-LOOPLU=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US1009916@cgn 1 1 354 @runat 26102004\_100225\_2092  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\*

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 414   | 43.6        | 34063  | 14    | US-10-114-170-96     |
| 2          | 413   | 43.5        | 48908  | 14    | US-10-114-170-137    |
| 3          | 385   | 40.5        | 558    | 15    | US-10-320-800-59     |
| 4          | 180   | 18.9        | 1067   | 17    | US-10-437-963-8605   |
| 5          | 179   | 18.8        | 542    | 10    | US-09-991-936-1494   |
| 6          | 179   | 18.8        | 623    | 16    | US-10-424-599-125718 |
| 7          | 175.5 | 18.5        | 998    | 17    | US-10-767-701-10509  |
| 8          | 174   | 18.3        | 582    | 9     | US-09-974-300-2101   |
| 9          | 173.5 | 18.3        | 957    | 10    | US-09-814-353-21657  |
| 10         | 172.5 | 18.2        | 673    | 16    | US-10-425-114-12447  |
| 11         | 171   | 18.0        | 459    | 9     | US-09-938-842A-1857  |
| 12         | 171   | 18.0        | 459    | 11    | US-09-938-842A-1857  |
| 13         | 171   | 18.0        | 636    | 9     | US-09-770-149-647    |
| 14         | 170   | 17.9        | 1004   | 16    | US-10-424-599-92887  |
| 15         | 168   | 17.7        | 657    | 9     | US-09-938-842A-1433  |
| 16         | 168   | 17.7        | 657    | 11    | US-09-938-842A-1433  |
| 17         | 167   | 17.6        | 816    | 16    | US-10-433-256-24     |
| 18         | 166.5 | 17.5        | 763    | 16    | US-10-425-114-7121   |
| 19         | 166.5 | 17.5        | 816    | 16    | US-10-425-114-31611  |
| 20         | 166.5 | 17.5        | 2523   | 10    | US-09-884-456-85     |
| 21         | 166.5 | 17.5        | 2523   | 10    | US-09-884-455-85     |
| 22         | 166   | 17.5        | 1156   | 17    | US-10-437-963-8609   |
| 23         | 165.5 | 17.4        | 459    | 18    | US-10-700-816-18     |
| 24         | 165.5 | 17.4        | 708    | 15    | US-10-272-459-37     |
| 25         | 165.5 | 17.4        | 874    | 16    | US-10-633-843-3      |
| 26         | 165.5 | 17.4        | 1056   | 15    | US-10-272-459-36     |
| 27         | 165.5 | 17.4        | 1148   | 15    | US-10-272-459-38     |
| 28         | 165.5 | 17.4        | 1956   | 15    | US-10-272-459-39     |
| 29         | 165.5 | 17.4        | 3297   | 9     | US-09-881-654-3      |
| 30         | 165.5 | 17.4        | 3297   | 16    | US-10-637-323-3      |
| 31         | 165.5 | 17.4        | 3297   | 17    | US-10-658-782-5      |
| 32         | 165   | 17.4        | 727    | 16    | US-10-425-114-3329   |
| 33         | 164.5 | 17.3        | 750    | 16    | US-10-425-114-26181  |
| 34         | 164.5 | 17.3        | 752    | 16    | US-10-425-114-2871   |
| 35         | 164.5 | 17.3        | 778    | 16    | US-10-425-114-25172  |
| 36         | 164.5 | 17.3        | 780    | 16    | US-10-425-114-26645  |
| 37         | 164.5 | 17.3        | 782    | 16    | US-10-425-114-16029  |
| 38         | 164.5 | 17.3        | 808    | 16    | US-10-425-114-23851  |
| 39         | 164.5 | 17.3        | 824    | 16    | US-10-425-114-13772  |
| 40         | 164.5 | 17.3        | 824    | 16    | US-10-425-114-13772  |
| 41         | 163.5 | 17.2        | 542    | 16    | US-10-242-535A-45794 |
| 42         | 163.5 | 17.2        | 542    | 16    | US-10-085-783A-45794 |
| 43         | 163.5 | 17.2        | 601    | 16    | US-10-388-934-236    |
| 44         | 163.5 | 17.2        | 729    | 16    | US-10-425-114-3924   |
| 45         | 163.5 | 17.2        | 771    | 16    | US-10-425-114-24323  |

ALIGNMENTS

RESULT 1  
US-10-114-170-96  
; Sequence 96, Application US/10114170  
; Publication No. US20030023075A1  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; Burland, Nicole T.  
; Perna, Nicole T.  
; Plunkett, Guy  
; Welch, Rod  
; TITLE OF INVENTION: No. US20030023075A1e1 Sequences of E. coli O157  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 1 South Pinckney Street  
; CITY: Madison  
; STATE: WI  
; COUNTRY: US  
; ZIP: 53701-2113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage





QY 37 AspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly 56  
Db 6902 AGTCTGACGGAAGAGTTCAGCATTCGAAATAACCATTCAGGAGACCCCTACGGT 6961  
QY 57 LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHis 76  
Db 6962 CTGCTGTTCACACCGCCCTTCATCTCTGTCTGAAGGCATTCATGGTTCATGTGCAC 7021  
QY 77 GluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGluAla 96  
Db 7022 GAAAAGGAATTCGGCCCGGCACTGAAAGACGGAAACCGGTGCGAGCATTATCGGCT 7081  
QY 97 HisGlyHisThrAspProAspLysThrGlyLysHisGluGlyProLeuGly--AsnGly 115  
Db 7082 GCGGCTCACATTGACCGGAAACACCGGCAACATCTTGCCCTCTGCTCCGGATGGA 7141  
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135  
Db 7142 CACCTGGCGACCTCCCTGCGCTCTCTGTGACGCATGACGGAAGGCACTACCCGCTC 7201  
QY 136 LeuAlaProArgLeu--ThrValLysGluIleLysGlyArgThrValMetIleHisAla 154  
Db 7202 CTGCCCCGAGAGTGAATCATTAAGAGAGATTAAAGGCGCTTCTCATGCTTCATGCT 7261  
QY 155 GlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174  
Db 7262 GCGGCTGATAACCATCATGACCATCCGAGCCCTCGGCGCTGTGTGCGAGATCGCC 7321  
QY 175 CysGlyValIle 178  
Db 7322 TGGCGCATCAT 7333

## RESULT 3

US-10-320-800-59  
; Sequence 59, Application US/10320800  
; Publication No. US20030215469A1  
; GENERAL INFORMATION:  
; APPLICANT: ROBINSON, ANDREW  
; APPLICANT: GORRINGE, ANDREW  
; APPLICANT: HUDSON, MICHAEL  
; APPLICANT: REDDIN, KAREN  
; TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE  
; FILE REFERENCE: 1581.0790001  
; CURRENT APPLICATION NUMBER: US/10/320,800  
; CURRENT FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: PCT/GB99/03626  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 59  
; LENGTH: 558  
; TYPE: DNA  
; ORGANISM: Neisseria meningitidis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(558)  
; OTHER INFORMATION: sodC CDS  
US-10-320-800-59

Alignment Scores:  
Pred. No.: 1,3e-39 Length: 558  
Score: 385.00 Matches: 77  
Percent Similarity: 64.71% Conservative: 22  
Best Local Similarity: 50.33% Mismatches: 52  
Query Match: 40.53% Indels: 2  
DB: 15 Gaps: 2

US-10-009-916A-1 (1-180) x US-10-320-800-59 (1-558)

QY 28 SerValThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThr 47  
Db 97 TCTATTGAAGTGAAGTGAACAACTTGATCCAGTAAACGGTAAACAAAGATGTGGGTACA 156

QY 48 ValThrPheThrAspThrAspLysGlyLeuGlnIleLysThrAspLeuLysGlyLeuPro 67  
Db 157 GTGACTATTACTGAATCTAACTATGGTCTTTGTGTTTACCCCTGATTTTACAAGGATTAAGC 216  
QY 68 AlaGlyGluHisGlyPheHisIleHisGlyGlySerCysGlyProAlaGluHisAsp 87  
Db 217 GAAGGCTTACATGGTTTCCACATCCATGAAACCCCAAGCTGTGAGCCCAAGAAAGAA 276  
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLys 107  
Db 277 GGTAAATGTAGCAGCTGGTTTAGCGCAGGCGGTCACTGGGATCCTAAAGGTGCAAAACAA 336  
QY 108 HisGlyGlyPro--LeuGlyAsnGlyHisLysGlyAspLeuProArgLeuValValLys 126  
Db 337 CATGTTTACCCATGGCAGAGTAGTGACACTTAGTGATTTTACCTGCACTAACTGTATTG 396  
QY 127 AlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThr--ValLysGluIle 145  
Db 397 CATGATGGCACAGCAACAAATCCTGTTTATAGCACCGCTCTTAAACATTTAGATGATGT 456  
QY 146 LysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrSerAspLysProLeuPro 165  
Db 457 CGCGGTCACCTCTATTATGATCCACACGGGTGTGATAATCACTCCGATCATCCAGTCCA 516  
QY 166 LeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
Db 517 CTGGCGGTGGCGGCCACCGATGCGCATGTGGCTGTGATT 555

## RESULT 4

US-10-437-963-8605  
; Sequence 8605, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 8605  
; LENGTH: 1067  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_15099C.1  
US-10-437-963-8605

Alignment Scores:  
Pred. No.: 6,19e-13 Length: 1067  
Score: 180.00 Matches: 52  
Percent Similarity: 45.71% Conservative: 28  
Best Local Similarity: 29.71% Mismatches: 57  
Query Match: 18.95% Indels: 38  
DB: 17 Gaps: 8

US-10-009-916A-1 (1-180) x US-10-437-963-8605 (1-1067)

QY 22 SerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAsnGlyIle 41  
Db 256 GGGCTGCTGTCGCCGACGCCACCAAGAGCGGTCGCCGTGCTC-----AAGGCAACC 309  
QY 42 LysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly--LeuGlnIleLys 60  
Db 310 TCCAGGTTGAGGGAGTTCGTCCCTCACCCAGGATGACCAAGGTCCTACACAGTGAAT 369

```
QY 61 ThrAspLeuLysGlyLeuProAlaGlyGluHisGlyHisIleHisGluGlyGly--- 79
Db 370 GTCCGTTGTACGGGACTTACTCTGGACTTACGGCTTCCACCTCCACGAGTTGGCGAT 429
QY 80 -----SerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGly 93
Db 430 ACTAGCAATGGTGCATATCAACAGCACCATTTTAACCCAAACAATTTGACG----- 483
QY 94 LeuGlnAlaHisGlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly 113
Db 484 -----CACGGTGCACACAGAGATGAAGTCCGTCATCGCGGTGACCTGGGA 528
QY 114 AsnGlyHisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGlu 133
Db 529 AAC-----ATGTTGCCAATGCTGAAGGTAGTACGAGGCA 564
QY 134 ThrLeuLeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrVal 150
Db 565 ACCATGTTGATAAGCAGATTCTCTGAGTGGCCCAAAATTCGTTGTTGGAGAGCATTC 624
QY 151 MetIleHis-----AlaGlyGlyAspAsnTyrSerAspLysPro 163
Db 625 GTTGTTCATGAGCTTGAAGATGATTGGGGAAGGGGTGCCCATGAGCTTAGT----- 675
QY 164 LeuProLeuGlyGlyGlyAlaAlaArgIleAlaCysGlyValIle 178
Db 676 CTCAGTACTGGAATGCTGGTGGGCGACTTGCAATGGGGTGTGTT 720
```

## RESULT 5

```
US-09-991-936-1494
; Sequence 1494, Application US/09991936
; Publication No. US20030073827A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1494
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Ctencecephalides felis
```

```
US-09-991-936-1494
```

```
Alignment Scores:
Pred. No.: 3,25e-13 Length: 542
Score: 179.00 Matches: 56
Percent Similarity: 46.33% Conservatives: 26
Best Local Similarity: 31.64% Mismatches: 47
Query Match: 18.84% Indels: 48
DB: 10 Gaps: 10
```

```
US-10-009-916A-1 (1-180) x US-09-991-936-1494 (1-542)
```

```
QY 20 IleThrSerValValLeuAlaCysSerValThrSerGluValHisMetIleAspAspAsn 39
Db 71 GTTACTAAAGCAGCTT-----TGC GTTTTAAACCGGGAAGTAAAG----- 109
QY 40 GlyIleLysGlnSerIleGlyThrValThrPheThrAspThr-----AspLysGlyLeu 57
Db 110 -----GGACCAATTACTTCGATCAAGCGGTCCAGAGCACCTGTC 151
```

```
QY 58 GlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
Db 152 ACACCTAACAGCATGGCTTAGTGGTTTAAAGCAAGGTGATCAGGTTTCCACATCCACGAA 211
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGly-----LeuGln 95
Db 212 TTCGGTGAC-----AGCACAATGGATGATTTTCAGCTGGGCCCACTTTTAAT 259
QY 96 AlaHisGly-----HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly 113
Db 260 CCCACGGTAAGACCATGGAGGACCTGATCTCTGATATCAGACATGTCCGGGACTTGGGA 319
QY 114 AsnGlyHisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGlu 133
Db 320 AAT-----CTTGTAGCTGATGCCGATGGAACGCTAAAGTG 355
QY 134 ThrLeuLeuAlaProArgLeuThrValLys-----GluIleLysGlyArgThrVal 150
Db 356 AAAATAACCGCAGTCAAATTTCTTACAAGGTCCTATGAGGTTATAGGCAGAACATGTG 415
QY 151 MetIleHisAlaGlyGlyAspAsnTyrSerAspLysProLeuProLeuGlyGly----- 168
Db 416 GTTGTACATGCTGATCCCGATGAT-----CTTGGATTAGGTGGTCATGAA 460
QY 169 -----GlyGlyAlaArgIleAlaCysGlyValIle 178
Db 461 CTTAGCAAGACCACTGGTAAATGCTGGAGCTGATTTGGCTTGTGGTATTG 511
```

## RESULT 6

```
US-10-424-599-125718
; Sequence 125718, Application US/10424599
; Publication No. US20040031072A1
```

## GENERAL INFORMATION:

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 125718
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Glycine max
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84532C.1
US-10-424-599-125718
```

```
Alignment Scores:
Pred. No.: 3,95e-13 Length: 623
Score: 179.00 Matches: 57
Percent Similarity: 50.32% Conservatives: 22
Best Local Similarity: 36.31% Mismatches: 52
Query Match: 18.84% Indels: 26
DB: 16 Gaps: 10
```

```
US-10-009-916A-1 (1-180) x US-10-424-599-125718 (1-623)
```

```
QY 35 MetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAsp 54
Db 81 ATCTCTAACCCCTGATGCTAAATTTCTGGTGTCTGTGTATCGTTAAGTTCAATCCACCTGTT 140
QY 55 LysGly-----LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPhe 73
Db 141 GGTGGCAAGGTCTCTGTCACGCTCCATCAAGGCTCTTAAGCTGTGTTAAACACGCTTTC 200
QY 74 HisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGly 93
Db 201 CACATCCACGAA-----TTCGGAACCTTGAAGGCTGAGGC 233
```

```

Qy 94 LeuGlnAlaHisGly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeu 112
Db 234 TGCAGAGACTGCTGGTGCCTCACTACAAACACACAAAGAGA---ACTCACGGTGGCCCCGAA 290
Qy 113 GlyAsn---GlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGly----- 129
Db 291 GATAATGAGACACATGTCGGTGATATGGGTAAATGTTGTGCCGACGAAACCGGTGAATCA 350
Qy 130 -----lleAlaLysGluThrLeuAlaProArgLeuThrValLys---GluIleLys 146
Db 351 GCTTTGATCTATGAGGATGCTTTACTC-----CAATTGCCGGTGAATTACGACTTCTT 404
Qy 147 GlyArgThrValMetIleHisAla-----GlyGlyAspAsnTyrSerAsp 161
Db 405 GGTAGATCTGTTGTTGCCACGCTGATGAAGATGACCTTGGCAGAGGAACCTTCGAAGAC 464
Qy 162 LysProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db 465 AGTAAG---ACTACTGGTCATGCTGGTGCCACACTTGCTTGTGGTGTGTCATC 512

RESULT 7
US-10-767-701-10509
; Sequence 10509, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated w
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 10509
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CIU517960_1
US-10-767-701-10509

```

```

RESULT 8
US-09-974-300-2101
; Sequence 2101, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2101
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2101

```

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS17960\_1  
US-10-767-701-10509



QY 106 GlyLys---HisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeu 123  
 Db 256 GGAACCAACCCATGGAGCTCTACTGCTGAAGACGTCATGCTGGTATTAGGCAATGTT 315  
 QY 124 ValValLysAlaAspGlyValLeuAlaLysGluThrLeuLeuAlaProArgLeuThrVal--- 142  
 Db 316 ATTGCTGCTGCCGATGAATGTCACAAAGTAGATATTACTGATAACAAATCGCCCTTACT 375  
 QY 143 -----LysGluLeuLysGlyArgThrValMetIleHisAlaGlyGlyAspAsnTyrSer 160  
 Db 376 GGACCAATCTGTTGTGACGTGCTTTAGTTGTTTATGCTGACCCGATGAT----- 429  
 QY 161 AspLysProLeuProLeuGlyGly-----GlyGlyVala 171  
 Db 430 -----TTGGTTTGGGAGGACATGATGTTGAGCAAAACCACTGGAATGCAGGAGCA 480  
 QY 172 ArgIleAlaCysGlyValIle 178  
 Db 481 CGTGTGTTGGTGTGTGATTT 501

## RESULT 11

US-09-938-842A-1857  
 ; Sequence 1857, Application US/09938842A  
 ; Patent No. US20020160378A1

## GENERAL INFORMATION:

; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SAME, AND METHODS OF USE

; CURRENT APPLICATION NUMBER: US/09/938,842A

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 1857

; LENGTH: 459

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-1857

Alignment Scores:  
 Pred. No.: 2,74e-12 Length: 459  
 Score: 171.00 Matches: 50  
 Percent Similarity: 45.83% Conservative: 16  
 Best Local Similarity: 34.72% Mismatches: 56  
 Query Match: 18.00% Indels: 22  
 DB: Gaps: 8

US-10-009-916A-1 (1-180) x US-09-938-842A-1857 (1-459)

QY 46 GlyThrValThrPheThrAspThrAspLysGlyLeu---GlnIleLysThrAspLeuLys 64  
 Db 46 GGGACTATCTTTTCCACCGAAGCGGATGGTGACCACTGTGAGTGAACAGATTCT 105  
 QY 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAla 84  
 Db 106 GGCCTTAAGCCTGGTCTTCATGTGTTCCATGTCTCTACTGCTCTTGGTGAC----- 153  
 QY 85 GluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAsp 103  
 Db 154 -----ACCCTAACCGTTGTCATGCTCTACTGCTCTTGGTGAC----- 198  
 QY 104 LysThrGlyLys---HisGluGlyProLeu---GlyAsnGlyHisLysGlyAspLeuPro 121  
 Db 199 -----GGTAAACACACACCGTCCCTGAGGATGCTAATCGACATGCTGTTGATCTAGGA 252

QY 122 ArgLeuValValLysAlaAspGlyValLeuAlaLysGluThrLeuLeu-----AlaPro 138  
 Db 253 AACATCACTGTTGGAGATGATGAACCTGCCACCTTCAATCACTGATTCGCCAGATTCT 312  
 QY 139 ArgLeuThrValLysGluLysGlyArgThrValMetIleHisAlaGlyGlyAspAsn 158  
 Db 313 CTCTACTGGACCAAACTCTATTGTTGGTAGGCTGTTGTTGTCATGCAGACCCCTGATGAC 372  
 QY 159 TyrSer-----AspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174  
 Db 373 CTCGGAAGGAGGCCCATGAATCAGCCTGGCTACTGGAACCGCAGCGCGCTGTTGCT 432  
 QY 175 CysGlyValIle 178  
 Db 433 TGGCGCATCATTT 444

## RESULT 12

US-09-938-842A-1857

; Sequence 1857, Application US/09938842A

; Publication No. US20040009476A9

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SAME, AND METHODS OF USE

; CURRENT APPLICATION NUMBER: US/09/938,842A

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 1857

; LENGTH: 459

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-1857

Alignment Scores:  
 Pred. No.: 2,74e-12 Length: 459  
 Score: 171.00 Matches: 50  
 Percent Similarity: 45.83% Conservative: 16  
 Best Local Similarity: 34.72% Mismatches: 56  
 Query Match: 18.00% Indels: 22  
 DB: Gaps: 8

US-10-009-916A-1 (1-180) x US-09-938-842A-1857 (1-459)

QY 46 GlyThrValThrPheThrAspThrAspLysGlyLeu---GlnIleLysThrAspLeuLys 64  
 Db 46 GGGACTATCTTTTCCACCGAAGCGGATGGTGACCACTGTGAGTGAACAGATTCT 105  
 QY 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAla 84  
 Db 106 GGCCTTAAGCCTGGTCTTCATGTGTTCCATGTCTCTACTGCTCTTGGTGAC----- 153  
 QY 85 GluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAsp 103  
 Db 154 -----ACCCTAACCGTTGTCATGCTCTACTGCTCTTGGTGAC----- 198  
 QY 104 LysThrGlyLys---HisGluGlyProLeu---GlyAsnGlyHisLysGlyAspLeuPro 121  
 Db 199 -----GGTAAACACACACCGTCCCTGAGGATGCTAATCGACATGCTGTTGATCTAGGA 252  
 QY 122 ArgLeuValValLysAlaAspGlyValLeuAlaLysGluThrLeuLeu-----AlaPro 138  
 Db 253 AACATCACTGTTGGAGATGATGAACCTGCCACCTTCAATCACTGATTCGCCAGATTCT 312

QY 139 ArgLeuThrValLysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsn 158  
DB 313 CTTACTGGACCAAACTATTGTTGGTAGGCTGTGTGTCCTGACACCTGATGAC 372  
QY 159 TyrSer-----AspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174  
DB 373 CTCGAAAGGGAGGCCATGAACTCAGCTGGCTACTGGAACGCGCGCGTGTGCT 432  
QY 175 CysGlyValIle 178  
DB 433 TCGGCGCATCAT 444

## RESULT 13

US-09-770-149-647/c

; Sequence 647, Application US/09770149

; Patent No. US20020059663A1

; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kricker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; FILE REFERENCE: 2024 (PARA-013PRV)

; CURRENT APPLICATION NUMBER: US/09/770,149

; PRIOR FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,506

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 647

; LENGTH: 636

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-770-149-647

Alignment Scores:  
Pred. No.: 4,318-12 Length: 636  
Score: 171.00 Matches: 50  
Percent Similarity: 45.83% Conservative: 16  
Best Local Similarity: 34.72% Mismatches: 56  
Query Match: 18.00% Indels: 22  
DB: 9 Gaps: 8

US-10-009-916A-1 (1-180) x US-09-770-149-647 (1-636)

QY 46 GlyThrValThrPheThrAspThrAspLysGlyLeu---GlnIleLysThrAspLeuLys 64  
DB 628 GGACATATCTTTTCCAGGAGGCGATGTTGTCACCATCTGTGAGTGGACAGTTTCT 569  
QY 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAla 84  
DB 568 GGCGCTTAAGCCTGCTCTTCATGGTTTCCATGTCCATGCTCTTGGTGAC----- 521  
QY 85 GluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAsp 103  
DB 520 -----ACCCTAACGGTTGCTGATGCTACTGTGTCACCATTTTCAACCCGAT 476

QY 104 LysThrGlyLys---HisGluGlyProLeu---GlyAsnGlyHisLysGlyAspLeuPro 121  
DB 475 -----GGTAAACACACACGGTGCCCTGAGGATGCTAATCGACATGCTGGTGATCTAGGA 422  
QY 122 ArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeuLeu-----AlaPro 138  
DB 421 AACATCACTGTTGGAGATGATGGAACCTGCCACCTTCAATCACTGATGTCAGATTCT 362  
QY 139 ArgLeuThrValLysGluIleLysGlyArgThrValMetIleHisAlaGlyGlyAspAsn 158  
DB 361 CTTACTGGACCAAACTCTATTGTTGGTAGGCTGTGTGTCCTGCAACGCGCGTGTGCT 302  
QY 159 TyrSer-----AspLysProLeuProLeuGlyGlyGlyAlaArgIleAla 174  
DB 301 CTCGAAAGGGAGGCCATGAACTCAGCTGGCTACTGGAACGCGCGCGTGTGCT 242  
QY 175 CysGlyValIle 178  
DB 241 TCGGCGCATCAT 230

## RESULT 14

US-10-424-599-92887/c

; Sequence 92887, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 92887

; LENGTH: 1004

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(1004)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_54892C.1

US-10-424-599-92887

Alignment Scores:  
Pred. No.: 1,098-11 Length: 1004  
Score: 170.00 Matches: 47  
Percent Similarity: 45.52% Conservative: 19  
Best Local Similarity: 32.41% Mismatches: 55  
Query Match: 17.89% Indels: 24  
DB: 16 Gaps: 8

US-10-009-916A-1 (1-180) x US-10-424-599-92887 (1-1004)

QY 46 GlyThrValThrPheThrAspThrAspLysGly---LeuGlnIleLysThrAspLeuLys 64  
DB 599 GGCGTGGCCACTCTCATCCAGAACGATGCGCCGACGACAGTTCCTGTCGCACTACT 540  
QY 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAla 84  
DB 539 GGCCCTTACTCCGGGGCTTCATGTTTTCACCTACATGATGATGTTGAT----- 492  
QY 85 GluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLys 104  
DB 491 ACGACAAATGGGTGATTTCTTACCTACATGATGATGATGATGATGATGATGATGAT 444  
QY 105 ThrGlyLysHisGluGlyProLeuGlyAsn---GlyHisLysGlyAspLeuProArgLeu 123  
DB 443 CTC---AACATGGTCTCCGGGAGATAAATTCGCCATCGCGGTGACCTGGGAAACATA 387

|    |     |  |     |
|----|-----|--|-----|
| QY | 17  | LeuThrSerIleThrSerValValLeuAlaCysSerVal-----ThrSerGluVal     | 33  |
| Db | 172 | TTGACAGTTGTTTCGGCGGCAAGAGCTGTTGCAGTGCTTAAAGGTACTTCTGATGC     | 231 |
|    |     |  |     |
| QY | 34  | HisMetIleAspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThr | 53  |
| Db | 232 | GAA-----GGAGTCTGTTACTTTGACCCAGAT                             | 258 |
| QY | 54  | AspLysGly---LeuGlnIleLysThrAspLeuLysGlyLeuProalaglyGluHisGly | 72  |
| Db | 259 | GACTCAGTCTCAACATGTGAATGTCGTATCACTGGTCTACTCCAGGGCCTCAATGA     | 318 |
|    |     |  |     |
| QY | 73  | PheHisIleHisGluGlyGly-----SerCysGlyProAlaGlu                 | 85  |
| Db | 319 | TTTCAATCTCCATGAGTTGGTGATACAACTAATGGATGATCTCAACAGACACCAATTC   | 378 |
|    |     |  |     |

```
Qy 86 HisAspGlyHisLeuThrAlaGlyLeuGlnAlaHisGlyHisTyrAspProAspLysThr 105
Db 379 AACCTTAACAATGACAC-----CACGGAGCTCCAGAGATGAG 417
Qy 106 GlyLysHisGluGlyProLeuGlyAsnGlyHisLysGlyAspLeuProArgLeuVal 125
Db 418 TGCGGTGATCGGGTGACCTGGGAAC-----ATAAATGCC 453
Qy 126 LysAlaAspGlyIleAlaLysGluThrLeuLeu-----AlaProArgLeuThrVal 142
Db 454 AATGCCGATGGCGTGGCAGAACCAATAGTGACACATCATCTCTGACCTGCTCT 513
Qy 143 LysGluIleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 514 AATCTGTGTGTGAAGAGCCCTTGTGGTTCACAGCTTAAGATGACCTCGGAAGGGT 573
Qy 156 GlyAspAsnTyrSerAspLysProLeuGlyGlyGlyAlaArgIleAlaCys 175
Db 574 GGCCATGAGCTTAGT-----CTGACCACTGGAAACGCGAGCGGAGATTGGCATGT 624
Qy 176 GlyVal 177
Db 625 GGTATG 630

RESULT 17
US-10-433-256-24
; Sequence 24, Application US/10433256
; Publication No. US20040081980A1
; GENERAL INFORMATION:
; APPLICANT: SANJANWALA, Madhusudan M.; YAO, Monique G.
; APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
; APPLICANT: ARVIZU, Chandra S.; RING, Ruijun Z.
; APPLICANT: LEE, Ernestine A.; DING, Li
; APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
; APPLICANT: YUE, Henry; TRIBOULRY, Catherine M.
; APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.
; APPLICANT: WARREN, Bridget A.; YANG, Junming
; APPLICANT: CHAWLA, Narinder K.; NGUYEN, Dannel B.
; APPLICANT: GANDHI, Ameena R.; LU, Yan
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; CURRENT APPLICATION NUMBER: US/10/433,256
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: PCT/US01/47429
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/254,308
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/256,189
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/257,713
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/262,706
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/266,020
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 55048919CB1
US-10-433-256-24
Alignment Scores:
Pred. No.: 1-98e-11 Length: 816
Score: 167.00 Matches: 46
Percent Similarity: 42.76% Conservative: 19
Best Local Similarity: 30.26% Mismatches: 53

Query Match: 17.58% Indels: 34
DB: 16 Gaps: 6
US-10-009-916A-1 (1-180) x US-10-433-256-24 (1-816)
Qy 45 IleGlyThrValThrPheThrAspLysGlyLeuGlnIleLysThrAspLeuLys 64
Db 113 ATCCGAGAGGTGACCTTGACCACGCGTGGCGTGTCTCTCGGTCTGACGCGTCC 172
Qy 65 GlyLeuProAlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyPro 83
Db 173 GGGCTGACGCTGGCTGGCAGCCCATGTTTCATGGAGTGGCGATTGTTCCGACCAA 232
Qy 84 -----AlaGluHisAspGlyHisLeuThrAlaGlyLeuGln 95
Db 233 GGCTTCCAGAAATCGGGCGGCATATCAACACGAGGATCACAAAGACA----- 280
Qy 96 AlaHisGlyHisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGlyAsnGly 115
Db 281 CCTCAGCGGCTTCTCAATCCG-----GAAGGGCCG----- 310
Qy 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGly----- 129
Db 311 GACTTTGGGAAATTGCCGATATATTGATGCGCGGAGCGCACCGTCAATGCCGAGCC 370
Qy 130 -----IleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGluIleLys 146
Db 371 TTCTCGGCACCTGTCTCGCTGGACGACGCGTCCTCCGCGCGAAGCTTCTCGATCCGAT 430
Qy 147 GlyArgThrValMetIleHisAlaGlyAspAsnTyrSerAspLysProLeuProLeu 166
Db 431 GGGTCCGCGCTGTATTTCACGCCGCCCCGATGACCATGTACGCAA-----CCGATC 484
Qy 167 GlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db 485 GCGCGGGCGGCGGCGCGGCTTGGTGTGCCGTATTC 520

RESULT 18
US-10-425-114-7121
; Sequence 7121, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7121
; LENGTH: 763
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700618619_FLI
US-10-425-114-7121
Alignment Scores:
Pred. No.: 2.09e-11 Length: 763
Score: 166.50 Matches: 53
Percent Similarity: 43.75% Conservative: 24
Best Local Similarity: 30.11% Mismatches: 70
Query Match: 17.53% Indels: 29
DB: 16 Gaps: 9
US-10-009-916A-1 (1-180) x US-10-425-114-7121 (1-763)
Qy 16 LeuLeuThrSerIleThrSerValValLeuAlaCysSerValThrSerGluValHisMet 35
```



```
Db 68 TTGCAGTTCAGATCGGTGATCACATAAACAATGCTGAAGCTGTGCTGCTT 127
QY 36 IleAspAsnGlyIleGlyThrValThrPheThrAspThrAspLys 55
Db 128 GGTAGCAGGAGGTGTCAG-----GGCACCATTCTTTCACCCAGAGGAGAT 178
QY 56 Gly---LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHis 74
Db 179 GGCCTTACCACCTGTACCCGGAAGTGTCTCTGGCTCAAGCTCGCTCCATGGTTCCAT 238
QY 75 IleHisGluGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeu 94
Db 239 GTGCATGCACTTGGTCAC-----ACCACCAATGGCTGCTGC 271
QY 95 GlnAlaHisGly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly 113
Db 272 ATGTCACTGGACCACTACATCTCT---GGAGCAAGGAACATGGACCAAGAT 328
QY 114 ---AsnGlyHisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLys 132
Db 329 GAGAACCGCATGCCGCTGATCTTGAATGTGCAGCTGGACAGATGGTGTGCAAC 388
QY 133 GluThrLeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThr 149
Db 389 ATTAATGTTACCGACAGCCAGATCCCACTGCTGGCCAACTCAATCATTTGCGAGCT 448
QY 150 ValMetIleHisAla-----GlyGlyAspAsnTyrSerAspLys 162
Db 449 GTTGTGTTACGCTGACCCCGATGATCTTGGAAAGGTGGACAGCTCAAGAGC 508
QY 163 ProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db 509 -----ACCGAAACGCTGCGCGCGTGTGCTTGTGGATCATC 547
```

## RESULT 19

```
US-10-425-114-31611
; Sequence 31611, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31611
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73194H01_FLI
US-10-425-114-31611
```

```
Alignment Scores:
Pred. No.: 2,3e-11 Length: 816
Score: 166.50 Matches: 49
Percent Similarity: 45.10% Conservative: 20
Best Local Similarity: 32.03% Mismatches: 55
Query Match: 17.53% Indels: 29
DB: 16 Gaps: 9
```

US-10-009-916A-1 (1-180) x US-10-425-114-31611 (1-816)

```
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly---Leu 57
Db 164 GATGGTGTCAAG-----GGCACCATTCTTTCACCCAGAGGAGATGCGCTTACC 214
```

```
QY 58 GlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
Db 215 ACTGTACCCGGAAGTGTCTCTGGCTCAAGCCTGGCTCCATGGGTTCATGTGCATGCA 274
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
Db 275 CTTGGTGCAC-----ACCACCAATGGCTGCTCAACT 307
QY 98 Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
Db 308 GGACCACTACTACATCTCT---GGAGCAAGGAACATGGAGCACCAGAGATGAGAACCGC 364
QY 116 HisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGluThrLeu 135
Db 365 CATGCCGCTGATCTTGGAAATGTGCAGCTGGAGCAGATGGTGTTCACCAACATTAATGT 424
QY 136 LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetIle 152
Db 425 ACCGACAGCAGATCCCACTGACTGGGCCAAACTCAATCATTTGCGAGAGCTGTGTGTT 484
QY 153 HisAla-----GlyGlyAspAsnTyrSerAspLysProLeuPro 165
Db 485 CAGCTGATCCCGATGATCTTGGAAAGGTGGACAGCTCAGGAAGAGCCCTC----- 538
QY 166 LeuGlyGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
Db 539 ---GGAAACGCTGCGCGCGTGTGCTTGTGGATCATC 574
```

## RESULT 20

```
US-09-884-456-85
; Sequence 85, Application US/09884456
; Publication No. US20030027317A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Ooi-Lim
; TITLE OF INVENTION: Hepatitis C virus protease
; FILE REFERENCE: 223002010005
; CURRENT APPLICATION NUMBER: US/09/884,456
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/253,230
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 08/709,177
; PRIOR FILING DATE: 1996-09-06
; PRIOR APPLICATION NUMBER: 08/440,548
; PRIOR FILING DATE: 1995-05-12
; PRIOR APPLICATION NUMBER: 08/350,884
; PRIOR FILING DATE: 1994-12-06
; PRIOR APPLICATION NUMBER: 07/680,296
; PRIOR FILING DATE: 1991-04-04
; PRIOR APPLICATION NUMBER: 07/505,433
; PRIOR FILING DATE: 1990-04-04
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vector cflSDp600
US-09-884-456-85
```

```
Alignment Scores:
Pred. No.: 1.1e-10 Length: 2523
Score: 166.50 Matches: 52
Percent Similarity: 47.85% Conservative: 26
Best Local Similarity: 31.90% Mismatches: 56
Query Match: 17.53% Indels: 29
DB: 10 Gaps: 9
```

US-10-009-916A-1 (1-180) x US-09-884-456-85 (1-2523)

```
Qy 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 7 ACAACCCCTGTTTCGGTTTGAAGGGTGACGGCCAGTTCAA-----GGTATTATTAAAC 60
Qy 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 61 TTCGAGCAGAAGAAAGTAATGGACCACTGAAGGTGTGGGAAGCATTTAAAGGACTGACT 120
Qy 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87
Db 121 GAAGCCCTGCATGATTCATGTTTCATGAGTTTGGAGAT----- 159
Qy 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106
Db 160 -----AATACAGCAGGCTGTACCACTCCAGTCTCTCACTTTAATCTT---CTATCCAGA 210
Qy 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125
Db 211 AAACACGGTGGCCCAAGGATGAAGAGAGGATGTTGGAGACTTTGGGCAATGTGACTGCT 270
Qy 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144
Db 271 GACAAAGATGTGTGGCCGATGCTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 330
Qy 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 331 CATTTGCATCATTTGGCCGCACACTGGTGTCCATGAAAGAGCAGATGACTTGGGCAAGGT 390
Qy 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175
Db 391 GGAATGAAGAAAGTACAAAG-----ACAGGAAACGCTGGAAGTCGTTTGGCTTGT 441
Qy 176 GlyValIle 178
Db 442 GGTGTAAATT 450
```

## RESULT 21

```
US-09-884-455-85
; Sequence 85, Application US/09884455
; Publication No. US20030064499A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Hepatitis C virus protease
; CURRENT APPLICATION NUMBER: US/09/884,455
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/253,675
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 08/709,177
; PRIOR FILING DATE: 1996-09-06
; PRIOR APPLICATION NUMBER: 08/440,548
; PRIOR FILING DATE: 1995-05-12
; PRIOR APPLICATION NUMBER: 08/350,884
; PRIOR FILING DATE: 1994-12-06
; PRIOR APPLICATION NUMBER: 07/680,296
; PRIOR FILING DATE: 1991-04-04
; PRIOR APPLICATION NUMBER: 07/505,433
; PRIOR FILING DATE: 1990-04-04
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vector cfISODp600
US-09-884-455-85
```

Alignment Scores: 1.1e-10 Length: 2523  
Pred. No.: 166.50 Matches: 52  
Score:

Percent Similarity: 47.85% Conservative: 26  
Best Local Similarity: 31.90% Mismatches: 56  
Query Match: 17.53% Indels: 29  
Db: 10 Gaps: 9

US-10-009-916A-1 (1-180) x US-09-884-455-85 (1-2523)

```
Qy 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 7 ACAACCCCTGTTTCGGTTTGAAGGGTGACGGCCAGTTCAA-----GGTATTATTAAAC 60
Qy 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 61 TTCGAGCAGAAGAAAGTAATGGACCACTGAAGGTGTGGGAAGCATTTAAAGGACTGACT 120
Qy 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87
Db 121 GAAGCCCTGCATGATTCATGTTTCATGAGTTTGGAGAT----- 159
Qy 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106
Db 160 -----AATACAGCAGGCTGTACCACTCCAGTCTCTCACTTTAATCTT---CTATCCAGA 210
Qy 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125
Db 211 AAACACGGTGGCCCAAGGATGAAGAGAGGATGTTGGAGACTTTGGGCAATGTGACTGCT 270
Qy 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144
Db 271 GACAAAGATGTGTGGCCGATGCTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 330
Qy 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 331 CATTTGCATCATTTGGCCGCACACTGGTGTCCATGAAAGAGCAGATGACTTGGGCAAGGT 390
Qy 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175
Db 391 GGAATGAAGAAAGTACAAAG-----ACAGGAAACGCTGGAAGTCGTTTGGCTTGT 441
Qy 176 GlyValIle 178
Db 442 GGTGTAAATT 450
```

## RESULT 22

```
US-10-437-963-8609
; Sequence 8609, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 8609
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15101C.1
US-10-437-963-8609
```

Alignment Scores: 4.32e-11 Length: 1156  
Pred. No.: 166.00 Matches: 53  
Score:

Percent Similarity: 46.29%      Conservative: 28  
Best Local Similarity: 30.29%      Mismatches: 55  
Query Match: 17.47%              Indels: 39  
DB: 17                              Gaps: 8

US-10-009-916A-1 (1-180) x US-10-437-963-8609 (1-1156)

```
QY 22 SerValValLeuAlaCysSerValThrSerGluValHisMetLeuAspAsnGlyIle 41
Db 361 GCGCTGTCGCGGACGCCACCAAGAGCGCGTCGCGTCTC-----AAGGGCACC 414
QY 42 LysGlnSerIleGlyThrValThrPheThrAspThrAspLysGlyLeuGlnIleLysTh 61
Db 415 TCCAGGTTGAGGAGTCGTCACCCCTCACCCAGATGACCAAGGTCTCTACACAGTGAAT 474
QY 61 xAspLeu---LysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlyGly-- 79
Db 475 GTCCGTGTGACGGGACTTACTCTGAGCTTCACGGCTTCACCTCCACAGTTGGCGA 534
QY 80 -----SerCysGlyProAlaGluHisAspGlyHisLeuThrAlaG1 93
Db 535 TACTACGAATGGGTGCATATCAACAGACCACATTTTAACCCCAACAAATTGACG---- 599
QY 93 yLeuGlnAlaHisGlyHisTyrrAspProAspLysThrGlyLysGluGlyProLeuG1 113
Db 590 -----CACGGTCACACAGAAAGATGAAGTCCGTCATCGCGGTGACCTGGG 633
QY 113 yAsnGlyHisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysG1 133
Db 634 AAC-----ATGTGTGCCAATGCTGAAGGTGATGACTGAGGC 669
QY 133 uThrLeuLeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrVa 150
Db 670 AACCAATGTTGATAGCAGATTCCTCTGAGTGGCCCAAAATTCGTGTTGGGAGAGCAAT 729
QY 150 lMetIleHis-----AlaGlyGlyAspAsnTyrrSerAspLysPr 163
Db 730 CGTTGTTCATGAGCTTGAAGATGATTGGGAGAGGTGGCCATGAGCTTAGT----- 781
QY 163 oLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyVal 177
Db 782 -CTCAGTACTGGAAGTCTGCTGGGCGACTTGCATGCGGTATG 823
```

## RESULT 23

```
US-10-700-816-18
; Sequence 18, Application US/10700816
; Publication No. US20040192629A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Zuoshang
; TITLE OF INVENTION: Allele-Specific RNA Interference
; FILE REFERENCE: UMY-038
; CURRENT APPLICATION NUMBER: US/10/700,816
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: 60/423,507
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/488,283
; PRIOR FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-700-816-18
```

Alignment Scores:  
Pred. No.: 1.39e-11      Length: 459  
Score: 165.50              Matches: 52  
Percent Similarity: 47.24%      Conservative: 25  
Best Local Similarity: 31.90%      Mismatches: 57  
Query Match: 17.42%              Indels: 29  
DB: 18                              Gaps: 9

US-10-009-916A-1 (1-180) x US-10-700-816-18 (1-459)

```
QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 4 ACGAAGCGCTGCTGCTGCTGAAGGGCGACGCCCGCAGTGCAG-----GCATCATCAAT 57
QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 58 TTCGAGCAGAGAAGAAAGTAATGACCAAGTGTGGGGAAGCATTAAGGACTGACT 117
QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87
Db 118 GAAGGCGCTGCATGGATTCCATGTTTCATGATTTGGAGAT----- 156
QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrrAspProAspLysThrGly 106
Db 157 -----AATACAGCAGGCTGTACCATGGCAGGTCTCATCTTAACTCT---CTATCCAGA 207
QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuVal 125
Db 208 AACACCGTGGGCCAAAGGATGAAGAGAGCATGTTGGAGACTTGGCAATGTGACTGCT 267
QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144
Db 268 GACAAAGATGGTGGCCGATGCTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 327
QY 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155
Db 328 CATTGCATCATTTGGCCGCACACTGGTGGTCCATGAAAAAGCAGATGACTTGGCAAGGT 387
QY 156 GlyAspAsnTyrrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175
Db 388 GGAATGAAGAAAGTACAAAG-----ACAGGAACGCTGGAAGTCTGTTGGCTGT 438
QY 176 GlyValIle 178
Db 439 GGTGTAAT 447
```

## RESULT 24

```
US-10-272-459-37
; Sequence 37, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-37
```

## Alignment Scores:

Pred. No.: 2.54e-11      Length: 708  
Score: 165.50              Matches: 52  
Percent Similarity: 47.24%      Conservative: 25  
Best Local Similarity: 31.90%      Mismatches: 57  
Query Match: 17.42%              Indels: 29  
DB: 15                              Gaps: 9

US-10-009-916A-1 (1-180) x US-10-272-459-37 (1-708)

```
QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 7 ACGAAGCGCTGTTGTGTTTGAAGGTGACGCCCGCAGTTCAA-----GGTATTATAAC 60
```

|    |     |  |     |
|----|-----|--|-----|
| Qy | 50  | PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro  | 67  |
|    |     | ::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::        |     |
| Db | 61  | TTCCAGCAGCAGGAAAGTAATGGACAGTGGAAGGTGTGGGGAAGCATTAAAGCACTGACT | 120 |
| Qy | 68  | AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp | 87  |
|    |     |  |     |
| Db | 121 | GAAGCCCTGCATGGATTCATGTTCCATGAGCTTGGAGAT-----                 | 159 |
| Qy | 88  | GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly | 106 |
|    |     | :::: :::: :::: :::: :::: :::: :::: ::::                      |     |
| Db | 160 | -----AATACAGCAGGCTGTACCAAGTCAGAGTCCTCACTTTAATCCT--CTATCCAGA  | 210 |
| Qy | 107 | LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal | 125 |
|    |     |  |     |
| Db | 211 | AAACACGGTGGGCCCAAGAGATGAAGAGAGGCATGTGGAGACTTGGCNAATGACTGCT   | 270 |
| Qy | 126 | LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- | 144 |
|    |     |  |     |
| Db | 271 | GACAAAGATGCTGTGGCCGATGCTTATTAAGATTTCTGTGATCTCACTCTCAGGAGAC   | 330 |
| Qy | 145 | -----IleLysGlyArgThrValMetIleHis-----AlaGly                  | 155 |
|    |     |  |     |
| Db | 331 | CATTGCATCAVTGGCCGCACACTGTGTGCTCCATGAAAAGCAGATGACTTGGCRAAAGT  | 390 |
| Qy | 156 | GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys    | 175 |
|    |     |  |     |
| Db | 391 | GGAATGAGAAAGATCAAAAG-----ACAGAAACGCTGGAAGTGGTTGGCTTGT        | 441 |
| Qy | 176 | GlyValIle  | 178 |
|    |     |  |     |
| Db | 442 | GGTGTAAATT   | 450 |
| Db |     |  |     |

Qy 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87  
 Db 185 GAAGGCCCTGCATGGATTCATGTTTCATGAGTTTGGAGAT----- 223  
 Qy 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly--HisTyrAspProAspLysThrGly 106  
 Db 224 -----AATACAGCAGGCTGTACAGTGCAGGCTCTACTTTAATCCCT---CTATCCAGA 274  
 Qy 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuValVal 125  
 Db 275 AAACACGGTGGCCCAAGAGTGAAGAGAGGCATGTGGAGACTTGGGCAATGTGACTGCT 334  
 Qy 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144  
 Db 335 GACAAAGATGCTGTGGCCGATGCTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 394  
 Qy 145 -----IleLysGlyArgThrValMetIleHis-----AlaGly 155  
 Db 395 CATTGCATCATTTGGCCGCACACTGGTGTCTCATGAAAGCAGATGACTTGGGCAAGGT 454  
 Qy 156 GlyAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175  
 Db 455 GGAATGAAGAAAGTACAAAG-----ACAGGAAACGCTGGAGTGGTGGCTGT 505  
 Qy 176 GlyValIle 178  
 Db 506 GGTGTAATT 514  
 RESULT 26  
 US-10-272-459-36  
 ; Sequence 36, Application US/10272459  
 ; Publication No. US20030124517A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PICHUANYES, Sergio  
 ; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
 ; TITLE OF INVENTION: PROTEINS AND USES THEREOF  
 ; FILE REFERENCE: PP17955.002 / 2301-17955  
 ; CURRENT APPLICATION NUMBER: US/10/272,459  
 ; CURRENT FILING DATE: 2002-10-15  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 36  
 ; LENGTH: 1056  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
 ; OTHER INFORMATION: of 38.8 kDa  
 US-10-272-459-36

US-10-009-916A-1 (1-180) x US-10-633-843-3 (1-874)

|    |     |  |     |
|----|-----|--|-----|
| Qy | 30  | ThrSerGluValHisMetIleAspAspAnglyIleLysGlnSerIleGlyThrValThr  | 49  |
|    |     |  |     |
|    |     |  |     |
|    |     |  |     |
| Db | 71  | ACGAAGGCGTGTGGTGTGTGAAGGCGAGCCCGCATGCAG                      | 124 |
|    |     |  |     |
|    |     |  |     |
|    |     |  |     |
| Qy | 50  | PheThrAspThrAspLysGlyLeuGlnIleLysThr                         | 67  |
|    |     |  |     |
|    |     |  |     |
|    |     |  |     |
| Db | 125 | ITCGAGCAGAAGAAAGTAATGACCACAGTCAAGCGTGTGGGAAGCAATTAAGACGTGACT | 184 |
|    |     |  |     |
|    |     |  |     |
|    |     |  |     |



```
Db 271 GACAAAGATGGTGGCGCATGCTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 330
Qy 145 -----lleYsGlyAqThrValMetIleHis:::
Db 331 CATTCGATCATTTGGCGGCACACTGGTGTCTCATGAAAGAGCAGATGCTGGGCAAGGT 390
Qy 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaAargIleAlaCys 175
Db 391 GGAATGAAGAAAGTACAAAG-----ACAGAAACGGTGGAGTCGTTGGCTGT 441
Qy 176 GlyValIle 178
Db 442 GGTGTAATT 450
RESULT 29
US-09-881-654-3
; Sequence 3, Application US/09881654
; Patent No. US20020146685A1
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARANGEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIEMENTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / PPI7039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/280,867
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1
; NAME/KEY: CDS
; LOCATION: (1)..(3297)
US-09-881-654-3
Alignment Scores:
Pred. No.: 2.14e-10 Length: 3297
Score: 165.50 Matches: 52
Percent Similarity: 47.24% Conservative: 25
Best Local Similarity: 31.90% Mismatches: 57
Query Match: 17.42% Indels: 29
DB: 9 Gaps: 9
US-10-009-916A-1 (1-180) x US-09-881-654-3 (1-3297)
Qy 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 7 ACAAGGCTGTTGTGTTTGAAGGGTGACGCCAGTTCAA-----GGTATTATTAAAC 60
Qy 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 61 TTCGAGCAGAGGAAAGTAATGACACAGTGAAGGTGTGGGAGCATTAAAGGACTGACT 120
Qy 68 AlaGlyGluHisGlyPheHisIleHisGlyGlySerCysGlyProAlaGluHisAsp 87
Db 121 GAAGCCCTGCATGGATTCCATGTTTCATGATTTGGAGAT-----GGTATTATTAAAC 159
Qy 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106
Db 160 -----AATACAGCAGGCTGTACACAGTCAGGTCTCTCACTTAACTCT---CTATCCAGA 210
```

```
Qy 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProAspGlyLeuValVal 125
Db 211 AAACACCGTGGCGCAAGGATGAAGAGGAGGATGTTGGAGACTTGGGCAATGTGACTGCT 270
Qy 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrValLysGlu--- 144
Db 271 GACAAAGATGGTGGCGCATGCTCTATTGAAGATTCTGTGATCTCACTCTCAGGAGAC 330
Qy 145 -----lleYsGlyAqThrValMetIleHis:::
Db 331 CATTCGATCATTTGGCGGCACACTGGTGTCTCATGAAAGAGCAGATGCTGGGCAAGGT 390
Qy 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaAargIleAlaCys 175
Db 391 GGAATGAAGAAAGTACAAAG-----ACAGAAACGGTGGAGTCGTTGGCTGT 441
Qy 176 GlyValIle 178
Db 442 GGTGTAATT 450
RESULT 30
US-10-637-323-3
; Sequence 3, Application US/10637323
; Publication No. US20040063092A1
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARANGEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIEMENTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / PPI7039.002
; CURRENT APPLICATION NUMBER: US/10/637,323
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/09/881,654
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/280,867
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1
; NAME/KEY: CDS
; LOCATION: (1)..(3297)
US-10-637-323-3
Alignment Scores:
Pred. No.: 2.14e-10 Length: 3297
Score: 165.50 Matches: 52
Percent Similarity: 47.24% Conservative: 25
Best Local Similarity: 31.90% Mismatches: 57
Query Match: 17.42% Indels: 29
DB: 16 Gaps: 9
US-10-009-916A-1 (1-180) x US-10-637-323-3 (1-3297)
Qy 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49
Db 7 ACAAGGCTGTTGTGTTTGAAGGGTGACGCCAGTTCAA-----GGTATTATTAAAC 60
Qy 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67
Db 160 -----AATACAGCAGGCTGTACACAGTCAGGTCTCTCACTTAACTCT---CTATCCAGA 210
```

Db 61 TTCGACGACGAAGAAAGTAATGGACCAAGTGAAGGTGTGGGGAAGCAATTAAGGACTGACT 120  
 QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87  
 Db 121 GAAGGCTCATGGATTCATGTTCTATGATTGGAGAT----- 159  
 QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106  
 Db 160 -----AATACAGCAGCGTGTACCAAGTGCAGGTCTCTCACTTTAATCCT---CTATCCAGA 210  
 QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuVal 125  
 Db 211 AAACACGGTGGGCCAAAGGATGAAGAGAGAGCATGTTGGAGACTTGGGCAATGTGCTGCT 270  
 QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrVal 144  
 Db 271 GACAAAGATGGTGGCCGATGTTCTATTGAAGATTCGTGATCTCACTCTCAGGAGAC 330  
 QY 145 -----IleLysGlyArgThrValMetIleHis----- 155  
 Db 331 CATTCGATCATTGGCCGCACACTGGTGTCCATGAAAGAGAGAGTGGGCAAGGT 390  
 QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175  
 Db 391 GGAATGAAAGAAAGTACAAAG-----ACAGGAAACGCTGGAAGTCGTTGGCTTGT 441  
 QY 176 GlyValIle 178  
 Db 442 GGTGTAATT 450

## RESULT 31

US-10-658-782-5  
 ; Sequence 5, Application US/10658782  
 ; Publication No. US20040142321A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ARCANGEL, Phillip  
 ; APPLICANT: CHIEN, David Y.  
 ; TITLE OF INVENTION: HCV ASSAY  
 ; FILE REFERENCE: 2300-19199  
 ; CURRENT APPLICATION NUMBER: US/10/658,782  
 ; PRIORITY FILING DATE: 2003-09-08  
 ; PRIOR APPLICATION NUMBER: 60/409,515  
 ; PRIOR FILING DATE: 2002-09-09  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 5  
 ; LENGTH: 3297  
 ; TYPE: DNA  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: MEFA 7.1 DNA sequence  
 US-10-658-782-5

Alignment Scores:  
 Pred. No.: 2,14e-10 Length: 3297  
 Score: 165.50 Matches: 52  
 Percent Similarity: 47.24% Conservative: 25  
 Best Local Similarity: 31.90% Mismatches: 57  
 Query Match: 17.42% Indels: 29  
 DB: 17 Gaps: 9

US-10-009-916A-1 (1-180) x US-10-658-782-5 (1-3297)

QY 30 ThrSerGluValHisMetIleAspAsnGlyIleLysGlnSerIleGlyThrValThr 49  
 Db 7 ACAAGCGCTGTTGTTGTTTGAAGGTGACGGCCAGTTCAA-----GGTATTATTAAAC 60  
 QY 50 PheThrAspThrAspLysGlyLeuGlnIleLysThr-----AspLeuLysGlyLeuPro 67  
 Db 61 TTCGACGACGAAGAAAGTAATGGACCAAGTGAAGGTGTGGGGAAGCAATTAAGGACTGACT 120  
 QY 68 AlaGlyGluHisGlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAsp 87

Db 121 GAAGGCTCATGGATTCATGTTCTATGATTGGAGAT----- 159  
 QY 88 GlyHisLeuThrAlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGly 106  
 Db 160 -----AATACAGCAGCGTGTACCAAGTGCAGGTCTCTCACTTTAATCCT---CTATCCAGA 210  
 QY 107 LysHisGluGlyProLeuGly---AsnGlyHisLysGlyAspLeuProArgLeuVal 125  
 Db 211 AAACACGGTGGGCCAAAGGATGAAGAGAGAGCATGTTGGAGACTTGGGCAATGTGCTGCT 270  
 QY 126 LysAlaAspGlyIleAlaLysGluThrLeuLeuAlaProArgLeuThrVal 144  
 Db 271 GACAAAGATGGTGGCCGATGTTCTATTGAAGATTCGTGATCTCACTCTCAGGAGAC 330  
 QY 145 -----IleLysGlyArgThrValMetIleHis----- 155  
 Db 331 CATTCGATCATTGGCCGCACACTGGTGTCCATGAAAGAGAGAGTGGGCAAGGT 390  
 QY 156 GlyAspAsnTyrSerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCys 175  
 Db 391 GGAATGAAAGAAAGTACAAAG-----ACAGGAAACGCTGGAAGTCGTTGGCTTGT 441  
 QY 176 GlyValIle 178  
 Db 442 GGTGTAATT 450

## RESULT 32

US-10-425-114-3329  
 ; Sequence 3329, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 3329  
 ; LENGTH: 727  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700260476\_FLI  
 US-10-425-114-3329

Alignment Scores:  
 Pred. No.: 3,05e-11 Length: 727  
 Score: 165.00 Matches: 52  
 Percent Similarity: 45.25% Conservative: 29  
 Best Local Similarity: 29.05% Mismatches: 68  
 Query Match: 17.37% Indels: 30  
 DB: 16 Gaps: 10

US-10-009-916A-1 (1-180) x US-10-425-114-3329 (1-727)

QY 14 IleSerLeuLeuThrSerIleThrSerValValLeuAlaCysSerValThrSerGluVal 33  
 Db 9 TTAACGACTTGAAGCGGTAGATGCTCATATACATAAACAAATGGTGAAGCTGTT 68  
 QY 34 HisMetIle---AspAspAsnGlyIleLysGlnSerIleGlyThrValThrPheThrAsp 52  
 Db 69 GCTGTGCTGTGGTAGCAGCGATGTTGTCAAG-----GGCACCATCTTTTTCACCCAA 119  
 QY 53 ThrAspLysGly---LeuGlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHis 71  
 Db 120 GTGGGAGATGGCCCTTACCACCTGTCCCGAAGTGTCTCTGGCTCAAGCTCGCTCCAT 179

```
Qy 72 GlyPheHisIleHisGluGlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThr 91
|||
Db 180 GGGTTTCATGTCATGCGCTGGTGAC-----ACCACC 212
|||
Qy 92 AlaGlyLeuGlnAlaHisGly---HisTyrAspProAspLysThrGlyLysHisGluGly 110
|||
Db 213 AATGGCTGCATGTCAACTGGACCACTACATCAATCT---GCCAGCAAGGACATGGAGCA 269
|||
Qy 111 ProLeuGly---AsnGlyHisGlyAspLeuProArgLeuValValLysAlaAspGly 129
|||
Db 270 CCAGAAGATGAGAACCGCCGCGGTGATCTTGGAAATGTGACAGCTGGAGCGAGATGGT 329
|||
Qy 130 IleAlaLysGluThrLeuLeuAlaProArgLeuThrVal-----LysGluIleLys 146
|||
Db 330 GTTGCAACATTAATGTTACCGACAGCCAGATCCCACTGACTGGGCCCAAACTCAATCATT 389
|||
Qy 147 GlyArgThrValMetIleHisAla-----GlyGlyAspAsnTyr 159
|||
Db 390 GGCAGAGCTGTGTGTTTCACGCTGACCCCGATGATCTTGGAAAGGGTGGACACGAGCTC 449
|||
Qy 160 SerAspLysProLeuProLeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
|||
Db 450 AGCAGAGC-----ACCGAAACGCTGGCGCGCTGTGCTTGTGGGATCATC 497
|||

RESULT 33
US-10-425-114-26181
; Sequence 26181, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26181
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4074-013-H7_FLI
US-10-425-114-26181

Alignment Scores:
Pred. No.: 3.69e-11 Length: 750
Score: 164.50 Matches: 49
Percent Similarity: 45.10% Conservative: 20
Best Local Similarity: 32.03% Mismatches: 55
Query Match: 17.32% Indels: 29
DB: 16 Gaps: 9

US-10-009-916A-1 (1-180) x US-10-425-114-26181 (1-750)
Qy 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly---Leu 57
|||
Db 87 GATGGTGTCAAG-----GGCACCATCTTTTTCACCCCAAGTGGGAGATGGCCCTACC 137
|||
Qy 58 GlnIleLysThrAspLeuLysGlyProAlaGlyGluHisGlyPheHisIleHisGlu 77
|||
Db 138 ACTGTACCGGAGAGTGTCTGTGCTCAAGCTGGCCCTCCATGGTGTTCATGTGTCATGCG 197
|||
Qy 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
|||
Db 198 CTTGGTGAC-----ACCACCAATGGTGCATGTCAACT 230
|||
Qy 98 Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
|||
```

```
Db 231 GGACACCACTACAATCTCT---GCCAGCAAGAAACATGAGACCCAGAGATGAGAACCGC 287
|||
Qy 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
|||
Db 288 CATCCCGGTGATCTTGGAAATGTGACAGCTGGACAGATGGTGTGTCAAACATTAATGTT 347
|||
Qy 136 LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetIle 152
|||
Db 348 ACCGACACCCAGATCCCACTGACTGGGCCAACTCAATCATTTGGCAGAGCTGTGTTGTT 407
|||
Qy 153 HisAla-----GlyGlyAspAsnTyrSerAspLysProLeuPro 165
|||
Db 408 CACGCTGACCCCGATGATCTTGGAAAGGGTGGACACGAGCTCAGCAAGAGC----- 458
|||
Qy 166 LeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178
|||
Db 459 ACCGAAACGCTGGCGCGCTGTGCTTGTGGGATCATC 497
|||

RESULT 34
US-10-425-114-2871
; Sequence 2871, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2871
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700236039_FLI
US-10-425-114-2871

Alignment Scores:
Pred. No.: 3.71e-11 Length: 752
Score: 164.50 Matches: 49
Percent Similarity: 45.10% Conservative: 20
Best Local Similarity: 32.03% Mismatches: 55
Query Match: 17.32% Indels: 29
DB: 16 Gaps: 9

US-10-009-916A-1 (1-180) x US-10-425-114-2871 (1-752)
Qy 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly---Leu 57
|||
Db 138 GATGGTGTCAAG-----GGCACCATCTTTTTCACCCCAAGTGGGAGATGGCCCTACC 188
|||
Qy 58 GlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
|||
Db 189 ACTGTACCGGAGAGTGTCTGTGCTCAAGCTGGCCCTCCATGGTGTTCATGTGTCATGCG 248
|||
Qy 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
|||
Db 249 CTTGGTGAC-----ACCACCAATGGTGCATGTCAACT 281
|||
Qy 98 Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
|||
Db 282 GGACACCACTACAATCTCT---GCCAGCAAGAAACATGAGACCCAGAGATGAGAACCGC 338
|||
Qy 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
|||
Db 339 CATCCCGGTGATCTTGGAAATGTGACAGCTGGAGCAGATGCTGTGTCAAACATTAATGTT 398
|||
```



```
QY 136 LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetile 152
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 399 ACCGACAGCCAGATCCCACTGACTGGGCCAAACTCAATCATTCGACAGAGCTGTGTGT 458
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 153 HisAla-----GlyGlyAspAsnTyrSerAspLysProLeuPro 165
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 459 CAGCTGACCCGATGATCTTGAAGGGTGGACAGCATCAGCAAGAGC----- 509
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 166 LeuGlyGlyGlyAlaAlaArgIleAlaCysGlyValile 178
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 510 ACCGGAAACGCTGGCGCGCTGTGTGCTTGTGGGATCATC 548
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
RESULT 35
US-10-425-114-25172
; Sequence 25172, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25172
; LENGTH: 778
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB732-001-H3_FLI
US-10-425-114-25172
Alignment Scores:
Pred. No.: 3,88e-11 Length: 778
Score: 164.50 Matches: 49
Percent Similarity: 45.10% Conservative: 20
Best Local Similarity: 32.03% Mismatches: 55
Query Match: 17.32% Indels: 29
DB: 16 Gaps: 9
US-10-009-916A-1 (1-180) x US-10-425-114-25172 (1-778)
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly---Leu 57
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 137 GATGGTGTCAAG-----GGCACCATCTTTTCCACCAAGAGGAGATGCCCTACC 187
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 58 GlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 188 GCTGTCACCGGAAGTCTCTGCGCTCAAGCTGGCTCCATGGGTTCATGTACATGCA 247
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 248 TTGTTGTGAC-----ACCACCAATGGATGTCATGTCAACT 280
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 98 Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 281 GGACCACTACATACTCT---GGAGCAGGAGCATGGGSCACCAAGATGAGAACCGC 337
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 338 CATGCTGGTATCTTGGAAACGTGACAGCTGGAGCAGATGGTGTCTAATATCAATGTC 397
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 136 LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetile 152
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 398 ACCGACAGCAGATCCCACTGACTGGGCCAAACTCAATCATTCGACAGAGCTGTGTGT 457
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 153 HisAla-----GlyGlyAspAsnTyrSerAspLysProLeuPro 165
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```
Db 458 CAGCTGATCTCTGATGATCTTGAAGGGTGGCCACGAGCTCAGCAAGAGC----- 508
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 166 LeuGlyGlyGlyAlaAlaArgIleAlaCysGlyValile 178
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 509 ACTGGAACCGCGGTGGCGCTGTGTGCTTGTGGGATCATTT 547
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
RESULT 36
US-10-425-114-26645
; Sequence 26645, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26645
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4605-009-C9_FLI
US-10-425-114-26645
Alignment Scores:
Pred. No.: 3,9e-11 Length: 780
Score: 164.50 Matches: 49
Percent Similarity: 45.10% Conservative: 20
Best Local Similarity: 32.03% Mismatches: 55
Query Match: 17.32% Indels: 29
DB: 16 Gaps: 9
US-10-009-916A-1 (1-180) x US-10-425-114-26645 (1-780)
QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly---Leu 57
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 117 GATGGTGTCAAG-----GGCACCATCTTTTCCACCAAGATGGGAGATGCCCTACC 167
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 58 GlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 168 ATGTCACCGGAAGTCTCTGCGCTCAAGCTGGCTCCATGGGTTCATGTGATGCG 227
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 228 CTTGTTGTAC-----ACCACCAATGGCTGCATGTCAACT 260
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 98 Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 261 GGACCACTACATACTCT---GGAGCAGGAAACATGAGCAGCAGCAGATGAGAACCGC 317
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 318 CATGCGGGTATCTTGGAAATGTGACAGCTGGAGCAGATGGTGTGCAACATTAATGTT 377
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 136 LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetile 152
      :::::  :::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 378 ACCGACAGCAGATCCCACTGACTGGGCCAAACTCAATCATTCGACAGAGCTGTGTGT 437
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 153 HisAla-----GlyGlyAspAsnTyrSerAspLysProLeuPro 165
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 438 CAGCTGACCCGATGATCTTGAAGGGTGGACAGCAGCTCAGCAAGAGC----- 488
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 166 LeuGlyGlyGlyAlaAlaArgIleAlaCysGlyValile 178
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 489 ACCGAAACGCTGGCGCGCTGTGTGCTTGTGGGATCATC 527
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

## RESULT 37

US-10-425-114-16029  
; Sequence 16029, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 16029

; LENGTH: 782

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3061-037-H2\_FLI

US-10-425-114-16029

## Alignment Scores:

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 3,91e-11 | Length:       | 782 |
| Score:                 | 164.50   | Matches:      | 49  |
| Percent Similarity:    | 45.10%   | Conservative: | 20  |
| Best Local Similarity: | 32.03%   | Mismatches:   | 55  |
| Query Match:           | 17.32%   | Indels:       | 29  |
| DB:                    | 16       | Gaps:         | 9   |

US-10-009-916A-1 (1-180) x US-10-425-114-16029 (1-782)

|    |     |  |        |
|----|-----|--|--------|
| Qy | 39  | AsnGlyLeuLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly---          | Leu 57 |
| Db | 149 | GATGGTGTCAAG-----GGCACCATCTTTTCAACCAAGTGGGAGATGGCCCTACC            | 199    |
| Qy | 58  | GlnIleLysThrAspLeuLysGlyLeuProAlaGluHisGlyGluHisGlyPheHisIleHisGlu | 77     |
| Db | 200 | ACTGTCCACCGAAGTGTCTGGCTCAAGCTGCCTCCATGGTTCATGTGTCATCGC             | 259    |
| Qy | 78  | GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis       | 97     |
| Db | 260 | CTTGGTGCAC-----ACCACCAATGGTGCATGTCAACT                             | 292    |
| Qy | 98  | Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly       | 115    |
| Db | 293 | GGACCACACTACAATCCT---GCGAGCAAGGAACATGGAGCACCAGAGATGAGAACC          | 349    |
| Qy | 116 | HisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGluThrLeu          | 135    |
| Db | 350 | CATCCGGTGATCTTGGAAATGTGACAGCTGGAGCAGATGGTGTGCAAACTTAATGTT          | 409    |
| Qy | 136 | LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetIle           | 152    |
| Db | 410 | ACCGACAGCCAGATCCCACTGCTGGGCCAACCAATCATTTGGCAGAGCTGTTGTTGTT         | 469    |
| Qy | 153 | HisIle-----GlyGlyAspAsnTyrSerAspLysProLeuPro                       | 165    |
| Db | 470 | CACGCTACCCCGATGATCTTGGAAAGGTGGACGAGCTCAGCAAGAGC-----               | 520    |
| Qy | 166 | LeuGlyGlyGlyAlaArgIleAlaCysGlyValIle                               | 178    |
| Db | 521 | ACCGAAACGCTGGCGCGGTGTGTGTGGGATCATC                                 | 559    |

## RESULT 38

US-10-425-114-25851  
; Sequence 25851, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 25851

; LENGTH: 808

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3960-014-F2\_FLI

US-10-425-114-25851

## Alignment Scores:

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 4,09e-11 | Length:       | 808 |
| Score:                 | 164.50   | Matches:      | 49  |
| Percent Similarity:    | 45.10%   | Conservative: | 20  |
| Best Local Similarity: | 32.03%   | Mismatches:   | 55  |
| Query Match:           | 17.32%   | Indels:       | 29  |
| DB:                    | 16       | Gaps:         | 9   |

US-10-009-916A-1 (1-180) x US-10-425-114-25851 (1-808)

|    |     |  |        |
|----|-----|--|--------|
| Qy | 39  | AsnGlyLeuLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly---          | Leu 57 |
| Db | 137 | GATGGTGTCAAG-----GGCACCATCTTTTCAACCAAGTGGGAGATGGCCCTACC            | 187    |
| Qy | 58  | GlnIleLysThrAspLeuLysGlyLeuProAlaGluHisGlyGluHisGlyPheHisIleHisGlu | 77     |
| Db | 188 | ACTGTCCACCGAAGTGTCTGGCTCAAGCTGCCTCCATGGTTCATGTGTCATCGC             | 247    |
| Qy | 78  | GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis       | 97     |
| Db | 248 | CTTGGTGCAC-----ACCACCAATGGTGCATGTCAACT                             | 280    |
| Qy | 98  | Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly       | 115    |
| Db | 281 | GGACCACACTACAATCCT---GCGAGCAAGGAACATGGAGCACCAGAGATGAGAACC          | 337    |
| Qy | 116 | HisLysGlyAspLeuProArgLeuValLysAlaAspGlyIleAlaLysGluThrLeu          | 135    |
| Db | 338 | CATCCGGTGATCTTGGAAATGTGACAGCTGGAGCAGATGGTGTGCAAACTTAATGTT          | 397    |
| Qy | 136 | LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetIle           | 152    |
| Db | 398 | ACCGACAGCCAGATCCCACTGCTGGGCCAACCAATCATTTGGCAGAGCTGTTGTTGTT         | 457    |
| Qy | 153 | HisIle-----GlyGlyAspAsnTyrSerAspLysProLeuPro                       | 165    |
| Db | 458 | CACGCTACCCCGATGATCTTGGAAAGGTGGACGAGCTCAGCAAGAGC-----               | 508    |
| Qy | 166 | LeuGlyGlyGlyAlaArgIleAlaCysGlyValIle                               | 178    |
| Db | 509 | ACCGAAACGCTGGCGCGGTGTGTGTGGGATCATC                                 | 547    |

## RESULT 39

US-10-425-114-13772  
; Sequence 13772, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 13772

; LENGTH: 810  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB143-041-B12\_FLI  
US-10-425-114-13772

Alignment Scores:  
Pred. No.: 4,11e-11 Length: 810  
Score: 164.50 Matches: 49  
Percent Similarity: 45.10% Conservative: 20  
Best Local Similarity: 32.03% Mismatches: 55  
Query Match: 17.32% Indels: 29  
DB: 16 Gaps: 9

US-10-009-916A-1 (1-180) x US-10-425-114-13772 (1-810)

QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly----Leu 57  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 147 GATGGTGTCAAG-----GGCACCATCTTTTTCACCCCAAGTGGGAGATGCCCTACC 197  
QY 58 GlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 198 ACTGTACCGGAAGTGTCTCTGCGCTCAAGCGCTGCTCATGGGTTTCATGTGCATGCG 257  
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 258 CITGGTGTAC-----ACCACCAATGGCTGCATGTCAACT 290  
QY 98 Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 291 GGACCACTACTACATCTCT--GGCAGCAAGAACATGGAGCACCAGAAAGATGAGAACCCG 347  
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 348 CATGCCGGTGTATCTTGGAAATGTGACAGCTGGAGCAGATGGTGTGCAAAACATTAATGTT 407  
QY 136 LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetIle 152  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 408 ACCGACAGCCAGATCCCACTGCTGGGCCAACTCAATCATTTGGCAGAGCTGTGTGTT 467  
QY 153 HisAla-----GlyGlyAspAsnTyrSerAspLysProLeuPro 165  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 468 CACGCTGACCCCGATGATCTTGGAAAGGTGGACACGAGCTCAGCAAGAGC----- 518  
QY 166 LeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 519 ACCGAAACGCTGGCGCGCGTGTGTGTTGTGGGATCATC 557

## RESULT 40

US-10-425-114-21202  
; Sequence 21202, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 21202

; LENGTH: 824  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3279-184-All\_FLI  
US-10-425-114-21202

Alignment Scores:  
Pred. No.: 4,21e-11 Length: 824  
Score: 164.50 Matches: 49  
Percent Similarity: 45.10% Conservative: 20  
Best Local Similarity: 32.03% Mismatches: 55  
Query Match: 17.32% Indels: 29  
DB: 16 Gaps: 9

US-10-009-916A-1 (1-180) x US-10-425-114-21202 (1-824)

QY 39 AsnGlyIleLysGlnSerIleGlyThrValThrPheThrAspThrAspLysGly----Leu 57  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 161 GATGGTGTCAAG-----GGCACCATCTTTTTCACCCCAAGTGGGAGATGCCCTACC 211  
QY 58 GlnIleLysThrAspLeuLysGlyLeuProAlaGlyGluHisGlyPheHisIleHisGlu 77  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 212 ACTGTACCGGAAGTGTCTCTGCGCTCAAGCGCTGCTCATGGGTTTCATGTGCATGCG 271  
QY 78 GlyGlySerCysGlyProAlaGluHisAspGlyHisLeuThrAlaGlyLeuGlnAlaHis 97  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 272 CITGGTGTAC-----ACCACCAATGGCTGCATGTCAACT 304  
QY 98 Gly---HisTyrAspProAspLysThrGlyLysHisGluGlyProLeuGly---AsnGly 115  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 305 GGACCACTACTACATCTCT--GGCAGCAAGAACATGGAGCACCAGAAAGATGAGAACCCG 361  
QY 116 HisLysGlyAspLeuProArgLeuValValLysAlaAspGlyIleAlaLysGluThrLeu 135  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 362 CATGCCGGTGTATCTTGGAAATGTGACAGCTGGAGCAGATGGTGTGCAAAACATTAATGTT 421  
QY 136 LeuAlaProArgLeuThrVal-----LysGluIleLysGlyArgThrValMetIle 152  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 422 ACCGACAGCCAGATCCCACTGCTGGGCCAACTCAATCATTTGGCAGAGCTGTGTGTT 481  
QY 153 HisAla-----GlyGlyAspAsnTyrSerAspLysProLeuPro 165  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 482 CACGCTGACCCCGATGATCTTGGAAAGGTGGACACGAGCTCAGCAAGAGC----- 532  
QY 166 LeuGlyGlyGlyAlaArgIleAlaCysGlyValIle 178  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 533 ACCGAAACGCTGGCGCGCGTGTGTGTTGTGGGATCATC 571

Search completed: November 1, 2004, 17:08:43

Job time : 1684 secs

***This Page Blank (uspto)***

| Result No. | Query |       |        | DB | ID                  | Description        |
|------------|-------|-------|--------|----|---------------------|--------------------|
|            | Score | Match | Length |    |                     |                    |
| 1          | 430   | 45.3  | 213    | 4  | US-09-328-353-6050  | Sequence 6050, App |
| 2          | 413.5 | 43.5  | 154    | 1  | US-07-641-346B-1    | Sequence 1, Appli  |
| 3          | 403.5 | 42.5  | 144    | 4  | US-09-489-039A-9511 | Sequence 9511, Ap  |
| 4          | 380.5 | 40.1  | 178    | 4  | US-09-543-681A-4466 | Sequence 4466, Ap  |
| 5          | 187.5 | 19.7  | 221    | 3  | US-08-679-493A-198  | Sequence 198, App  |
| 6          | 185   | 19    | 218    | 3  | US-08-679-493A-195  | Sequence 195, App  |
| 7          | 179.5 | 18.9  | 217    | 3  | US-08-679-493A-196  | Sequence 196, App  |
| 8          | 178   | 18.7  | 195    | 3  | US-08-679-493A-189  | Sequence 189, App  |
| 9          | 178   | 18.7  | 202    | 3  | US-08-679-493A-197  | Sequence 197, App  |
| 10         | 176.5 | 18.6  | 150    | 2  | US-08-722-050-11    | Sequence 11, Appl  |
| 11         | 176.5 | 18.6  | 150    | 4  | US-09-883-985-11    | Sequence 11, Appl  |
| 12         | 175.5 | 18.5  | 150    | 3  | US-08-679-493A-199  | Sequence 199, Appl |
| 13         | 175.5 | 18.5  | 166    | 3  | US-08-679-493A-209  | Sequence 209, App  |
| 14         | 172.5 | 18.2  | 151    | 2  | US-08-722-050-5     | Sequence 5, Appli  |
| 15         | 172.5 | 18.2  | 151    | 4  | US-09-883-985-5     | Sequence 5, Appli  |
| 16         | 170.5 | 17.9  | 152    | 6  | 5171680-3           | Patent No. 5171680 |
| 17         | 168   | 17.7  | 152    | 3  | US-08-679-493A-203  | Sequence 203, App  |
| 18         | 168   | 17.7  | 153    | 2  | US-08-722-050-3     | Sequence 3, Appli  |
| 19         | 168   | 17.7  | 153    | 4  | US-09-883-985-3     | Sequence 3, Appli  |
| 20         | 167   | 17.6  | 152    | 3  | US-08-679-493A-208  | Sequence 208, App  |
| 21         | 166.5 | 17.5  | 150    | 2  | US-08-722-050-4     | Sequence 4, Appli  |
| 22         | 166.5 | 17.5  | 150    | 4  | US-09-883-985-4     | Sequence 4, Appli  |
| 23         | 166.5 | 17.5  | 151    | 2  | US-08-722-050-6     | Sequence 6, Appli  |
| 24         | 166.5 | 17.5  | 151    | 3  | US-09-202-832-16    | Sequence 16, Appl  |
| 25         | 166.5 | 17.5  | 151    | 3  | US-08-679-493A-205  | Sequence 205, App  |
| 26         | 166.5 | 17.5  | 151    | 4  | US-09-883-985-6     | Sequence 6, Appli  |
| 27         | 166.5 | 17.5  | 201    | 3  | US-08-679-493A-190  | Sequence 190, App  |

STREET: 1815 NORTH UNIVERSITY STREET  
CITY: PEORIA  
STATE: IL  
COUNTRY: U.S.A.  
ZIP: 61604  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE - 5.25 INCH, 360 Kb STORAGE  
COMPUTER: IBM XT COMPATIBLE  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: MULTIMATE ADVANTAGE II  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/641.346B  
FILING DATE: 19910116  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: N/A  
APPLICATION NUMBER: N/A  
FILING DATE: N/A  
ATTORNEY/AGENT INFORMATION:  
NAME: RIBANDO, CURTIS P.  
REGISTRATION NUMBER: 27,976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: FTS 360-4513, COM 309/685-4011, x513  
TELEFAX: COM 309/685-4128  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 154  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-641-346B-1

Query Match 43.5%; Score 413.5; DB 1; Length 154;  
Best Local Similarity 52.0%; Pred. No. 2.2e-40;  
Matches 79; Conservative 19; Mismatches 53; Indels 1; Gaps 1;  
QY 28 SVTSEVMIDNGIKQSGTGTFTTDDKGLQIKTLKGLPAGHGHHGSGCGPAEHD 87  
DB 2 STTVKALPTGPGKEVGTWISAPGGLHFKVNMKLTGPGYHGFVHNPSCAPGEKD 61  
QY 88 GHLTAGLAQGHYDPDKTKEGHPGNGHKGDLPELVVKADGIAKETLLAPRL-TVKEIK 146  
DB 62 GKIVPALAAGGHYDPGNTHHLPBGDGHMGDLPRLSANADGKVSFTVVAHLKKLAIEIK 121  
QY 147 GRTVMHAGGNNYSKPLPLGGGGGARIACGVI 178  
DB 122 QRLSLMVHVGGNNYSKPKPEPLGGGGARFACGVI 153

RESULT 3  
US-09-489-039A-9511  
; Sequence 9511, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489, 039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9511  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9511

Query Match 42.5%; Score 403.5; DB 4; Length 144;  
Best Local Similarity 58.7%; Pred. No. 2.9e-39;  
Matches 74; Conservative 15; Mismatches 36; Indels 1; Gaps 1;

QY 54 DKGLQIKTDLKGLPAGHGHHGSGCGPAEHDGHLTAGLAQGHYDPDKTKGHEGPIG 113  
DB 18 DRGLEFAPTLRALPGKHGFHHAEGSCQPAKMGKAVAGAAGGHYDPOHTGKHEGPIG 77  
QY 114 NGHKGDLPLRVVKADGIAKETLLAPRL-TVKEIKGRTVMHAGGNNYSKPLPLGGGGAR 172  
DB 78 AGHLGDLPLRVVNDAGVADQPIIAPRLKTLAIVKGLMVHVGNDNVADSPQLGGGGAR 137  
QY 173 IACGVI 178  
DB 138 PACGVI 143

RESULT 4  
US-09-543-681A-4466  
; Sequence 4466, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 4466  
; LENGTH: 178  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-4466

Query Match 40.1%; Score 380.5; DB 4; Length 178;  
Best Local Similarity 50.0%; Pred. No. 2e-36;  
Matches 85; Conservative 13; Mismatches 67; Indels 5; Gaps 3;  
QY 14 ISLITS---ITSVVLACSVTSEVMIDNGIKQSGTGTFTTDDKGLQIKTLKGLPAGE 70  
DB 8 IPLLLSGLLFTSVASAASLDVTLKEALPTGAGDDIGVVVTITETDYGLLFTPKLTGLTFCV 67  
QY 71 HGFTHIEGSCGPAEHDGHLTAGLAQGHYDPDKTKEGHPG-NGHKGDLPLRVVKADG 129  
DB 68 HGFTHIANGSCPEPMKDGKVPALKAGHLDPENKGVHLGPNKKGHLDLPLVANSKG 127  
QY 130 IAKETLLAPRLT-VKEIKGRTVMHAGGNNYSKPLPLGGGGARFACGVI 178  
DB 128 DADYAVLAPRLTKLDQIKDKALMVHVGGNNYSNPEALGGGGARFACGVI 177

RESULT 5  
US-08-679-493A-198  
; Sequence 198, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Ethan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95  
; CURRENT APPLICATION NUMBER: US/08/679,493A  
; CURRENT FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 60/001203  
; PRIOR FILING DATE: 1995-07-14  
; PRIOR APPLICATION NUMBER: 60/003,112  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 198  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: spinach  
US-08-679-493A-198

Query Match 19.7%; Score 187.5; DB 3; Length 221;

```

Best Local Similarity 30.7%; Pred. No. 1.1e-13;
Matches 55; Conservative 24; Mismatches 73; Indels 27; Gaps 8;

QY 13 TISLLTSITSVVLACSVTSEVHMIDDNGIKQSIGCTVFTTDDKG-LQIKTDLKLPGAGEH 71
   :||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 52 SLSLSTSAAKPLTIVAATKKAVALKGTSNVEGVTLTQEDDGPPTVNRVISGLAPGKH 111
   :||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 72 GFHIHGGCGPAEHDGHLTAGLQAHG-HYDPDKTKGHEGPLN-GHKGDLPRLWVKADG 129
   ||||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 112 GFHLHEFGT-----TNGCMSTGPHFNPDKK-THGAPEDVVRHAGDLGNIVANTDG 160
   ||||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 130 IAKETLL---APELTVKEIKGRVTMHH-----AGDNYSDKPLPLGGGGARIACGVI 178
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 161 VABATIVDNIQPLTGPNSVVGRALVWHELEDLKGKGHELSP---TTGNAGGRLACGVW 216
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 6
US-08-679-493A-195
; Sequence 195, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 195
; LENGTH: 218
; TYPE: PRT
; ORGANISM: petunia
US-08-679-493A-195

Query Match 19.5%; Score 185; DB 3; Length 218;
Best Local Similarity 29.2%; Pred. No. 2.1e-13;
Matches 56; Conservative 32; Mismatches 74; Indels 30; Gaps 9;

QY 1 MKIKLFFVTSIVTISLTSITSVLACSVTSEVHMIDDNGIKQSIGTFTTDDKG-LQI 59
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 38 VSLKVKSKTQSLTSSVTSKPFIVFAATKKAVAL--KGTSNVEGVVTLTQDDGPTTV 95
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 60 KTDLKLGPAGEGHFHHEGSCGPAEHDGHLTAGLQAHG-HYDPDKTKGHEGPLN--GH 116
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 96 KVRITGLAPLGHGLHLEHFGD-----TTNGCMSTGPHFNP--GLTHGAPGDEVHR 144
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 117 KGDLPLRVKADGIAKETLLAPRLTV---KEIKGRVTMHH-----AGDNYSDKPLPL 166
   ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 145 AGDLGNIEANASVAATLVNDQNIPLSGNSVVGRLVWHELEDLKGKGHELSP---LTT 201
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 167 GCGGARIAACGVI 178
   ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 202 GNAGGRLACGVW 213
   ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 7
US-08-679-493A-196
; Sequence 196, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01

```





```
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 150 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-883-985-11

Query Match          18.5%; Score 176.5; DB 4; Length 150;
Best Local Similarity 33.8%; Pred. No. 1.2e-12;
Matches 50; Conservative 18; Mismatches 49; Indels 31; Gaps 7;

QY 46 GTVTFTDQKG-LQIKTDLKLPGAGHGFHIEGGSCGPAEHGHLTAG-----LQAHGHY 100
Db 15 GVVRFQDDGDVTVEGKISGLDGNHGFHIVFGD-----NTNGCLSGAGHFPNQKNHG 70
QY 101 DPDKTKHEGPLNGHKGDLPRLVVKADGIKATETLLAPRLTVK---EIKGRTVMIH---- 153
Db 71 SPKDADRHVGD LGN-----VTAEGGVAQFNFTDPQISLKGERSIIGRTAVVHEKQD 121
QY 154 ---AGGDNYSKPLPLGGGGARIACGVI 178
Db 122 DLKGGD---DESLKTNAGGRLACGVI 146

RESULT 12
US-08-679-493A-199
/ Sequence 199, Application US/08679493A
/ Patent No. 6303295
/ GENERAL INFORMATION:
/ APPLICANT: Taylor, Ethan W.
/ TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
/ FILE REFERENCE: 55-95
/ CURRENT APPLICATION NUMBER: US/08/679,493A
/ CURRENT FILING DATE: 1996-07-12
/ PRIOR APPLICATION NUMBER: 60/001203
/ PRIOR FILING DATE: 1995-07-14
/ PRIOR APPLICATION NUMBER: 60/003,112
/ PRIOR FILING DATE: 1995-09-01
/ NUMBER OF SEQ ID NOS: 216
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 199
/ LENGTH: 150
/ TYPE: PRT
/ ORGANISM: Xenopus laevis
US-08-679-493A-199

Query Match          18.5%; Score 175.5; DB 3; Length 150;
Best Local Similarity 33.8%; Pred. No. 1.6e-12;
Matches 50; Conservative 18; Mismatches 49; Indels 31; Gaps 7;

QY 46 GTVTFTDQKG-LQIKTDLKLPGAGHGFHIEGGSCGPAEHGHLTAG-----LQAHGHY 100
Db 15 GVVRFQDDGDVTVEGKISGLDGNHGFHIVFGD-----NTNGCLSGAGHFPNQKNHG 70
QY 101 DPDKTKHEGPLNGHKGDLPRLVVKADGIKATETLLAPRLTVK---EIKGRTVMIH---- 153
Db 71 SPKDADRHVGD LGN-----VTAEGGVAQFKFTDPQISLKGERSIIGRTAVVHEKQD 121
QY 154 ---AGGDNYSKPLPLGGGGARIACGVI 178
Db 122 DLKGGD---DESLKTNAGGRLACGVI 146

RESULT 13
US-08-679-493A-209
/ Sequence 209, Application US/08679493A
/ Patent No. 6303295
/ GENERAL INFORMATION:
/ APPLICANT: Taylor, Ethan W.
/ TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
```

```
/ FILE REFERENCE: 55-95
/ CURRENT APPLICATION NUMBER: US/08/679,493A
/ CURRENT FILING DATE: 1996-07-12
/ PRIOR APPLICATION NUMBER: 60/001203
/ PRIOR FILING DATE: 1995-07-14
/ PRIOR APPLICATION NUMBER: 60/003,112
/ PRIOR FILING DATE: 1995-09-01
/ NUMBER OF SEQ ID NOS: 216
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 209
/ LENGTH: 166
/ TYPE: PRT
/ ORGANISM: loggerhead
US-08-679-493A-209

Query Match          18.5%; Score 175.5; DB 3; Length 166;
Best Local Similarity 33.1%; Pred. No. 1.8e-12;
Matches 50; Conservative 31; Mismatches 45; Indels 25; Gaps 11;

QY 41 IKQSI-GTVTFTDQKG-LQIKTDLKLPGAGHGFHIEGGSCGPAEHGHLTAGLQAHG 98
Db 23 VKEPVKGLIYFEOQNGPVTLSGSIITGLTEGKHGPHVHEFGD---NTNGCTTSAG----A 74
QY 99 HYDPDKTKG-HEGPLGN-GHKGLDPLRVVKADGIA---KETLLAPRLT-VKSIKGRVTM 151
Db 75 HFNP--PGKNHGGPQDNERHVGDLGNVIANKGVABVCIKDSLIS--LTGSQSIIGRTMV 130
QY 152 IHAGGDNY----SDKPLPLGGGGARIACGVI 178
Db 131 VHEKEDDLGKGNDKSLKTNAGSRLACGVI 161

RESULT 14
US-08-722-050-5
/ Sequence 5, Application US/08722050
/ Patent No. 5871729
/ GENERAL INFORMATION:
/ APPLICANT: YU, GUO-LIANG
/ APPLICANT: ROSEN, CRAIG A.
/ APPLICANT: FRASER, CLAIRE M.
/ APPLICANT: GOCAYNE, JEANNINE D.
/ TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
/ NUMBER OF SEQUENCES: 16
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
/ STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
/ CITY: WASHINGTON
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/722,050
/ FILING DATE: 23-JAN-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/225,757
/ FILING DATE: 11-APR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: STEFFE, ERIC K.
/ REGISTRATION NUMBER: 36,688
/ REFERENCE/DOCKET NUMBER: 1488.1020001/EKS/AJK
/ TELEPHONE: (202) 371-2600
/ TELEFAX: (202) 371-2540
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 151 amino acids
/ TYPE: amino acid
```





## RESULT 20

US-08-679-493A-208  
; Sequence 208, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Ethan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95  
; CURRENT APPLICATION NUMBER: US/08/679,493A  
; CURRENT FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 60/001203  
; PRIOR FILING DATE: 1995-07-14  
; PRIOR APPLICATION NUMBER: 60/003,112  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 208  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: blueshark  
US-08-679-493A-208

Query Match 17.6%; Score 167; DB 3; Length 152;  
Best Local Similarity 32.5%; Pred. No. 1.6e-11;  
Matches 49; Conservative 26; Mismatches 52; Indels 24; Gaps 8;

QY 40 GIKQSIGVTFTDIDKGLQIYKTDLKGHPAGHGFHIEHGGSCGPAEHDGHLTAGLQAHG 98  
DB 9 GTGEVTGLVFEQAADGPTLKGSTITGLTPGKHGFHVAFGD----NTNGCISAG----P 60  
QY 99 HYDPDKTGKHEGFLG-NGHKGDPLRLVVKADGIKAKETLLAPRLTV---KEIKGRWTMIIH- 153  
DB 61 HYNP-FSKNHGCGPDEERHVGDLGNVEANGVGAFFKIDRQLHLSGERSIIGRTLVVHE 119  
QY 154 -----AGDNYSDKPLPLGGGGARIACGVI 178  
DB 120 KEDDLGKGD---EESLRTGNAGSLACGVI 147

## RESULT 21

US-08-722-050-4  
; Sequence 4, Application US/08722050  
; Patent No. 5871729  
; GENERAL INFORMATION:  
; APPLICANT: YU, GUO-LIANG  
; APPLICANT: ROSEN, CRAIG A.  
; APPLICANT: FRASER, CLAIRE M.  
; APPLICANT: GOCAYNE, JEANNINE D.  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/722,050  
; FILING DATE: 23-JAN-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/225,757  
; FILING DATE: 11-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.1020001/EKS/AJK

## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 150 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-722-050-4

Query Match 17.5%; Score 166.5; DB 2; Length 150;  
Best Local Similarity 29.3%; Pred. No. 1.8e-11;  
Matches 49; Conservative 27; Mismatches 52; Indels 39; Gaps 8;

QY 30 TSEVHMDNGIKQSIGVTFTDIDKGLQIYKTDLKGHPAGHGFHIEHGG-----SCG 82  
DB 1 TKAVCVLKGDPVQ--GTIHFEAKGDTVVVTGSGITGLTEGHDGPHVHQFGONTQCTSAG 58  
QY 83 PAEHDGHLTAGLQAHDGPDYDKTGKHEGFLG-NGHKGDPLRLVVKADGIKAKETLLAPRLT 141  
DB 59 P-----HENP-LSKKHGKPKDEERHVGDLGNVTADKNGVAIVDIDVPLIS 102  
QY 142 VK---EIKGRWTMIIH-----AGDNYSDKPLPLGGGGARIACGVI 178  
DB 103 LSGEYSIIIGRTWVHEKPDLDLGRGNEESTK---TGNAGSELACGVI 146

## RESULT 22

US-09-883-985-4  
; Sequence 4, Application US/09883985  
; Patent No. 6635252  
; GENERAL INFORMATION:  
; APPLICANT: YU, GUO-LIANG  
; ROSEN, CRAIG A.  
; FRASER, CLAIRE M.  
; GOCAYNE, JEANNINE D.  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/883,985  
; FILING DATE: 20-Jun-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/203,607  
; FILING DATE: 02-DEC-1998  
; APPLICATION NUMBER: US 08/722,050  
; FILING DATE: 23-JAN-1997  
; APPLICATION NUMBER: US 08/225,757  
; FILING DATE: 11-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.1020003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 150 amino acids

```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-883-985-4

Query Match      17.5%; Score 166.5; DB 4; Length 150;
Best Local Similarity 29.3%; Pred. No. 1.8e-11;
Matches 49; Conservative 27; Mismatches 52; Indels 39; Gaps 8;

QY 30 TSEVHMIDNGIKQSIGTFTTDTKGLPAGHGFIHEGSGCGPAEHDGHLTAGLQAHGYDPD 103
Db 1 TRAVCVLKGDPVQ--GTTHFEAKGDTVVVVTGSLTGLTEGDHGFHVHQFGDNTQGCTSA 58
QY 83 PRAEHDGHLTAGLQAHGYDPDKTGKHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLT 141
Db 59 P-----HFNP-LSKKGGPKDERHVGDLGNVTADKXGVAIVDIVDPLIS 102
QY 142 VK---EIKGRTVMIH-----AGGDNYSKPLPLGGGGGARIACGVI 178
Db 103 LSGEYSIIIGRTVMVHEKPDGLGRGNEESTK---TGNAGSRLACGVI 146

RESULT 23
US-08-722-050-6
; Sequence 6, Application US/08722050
; Patent No. 5871729
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: FRASER, CLAIRE M.
; APPLICANT: GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,050
; FILING DATE: 23-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020001/EKS/AJK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-722-050-6

Query Match      17.5%; Score 166.5; DB 2; Length 151;
Best Local Similarity 34.7%; Pred. No. 1.8e-11;
Matches 50; Conservative 19; Mismatches 54; Indels 21; Gaps 8;

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-883-985-4

Query Match      17.5%; Score 166.5; DB 3; Length 151;
Best Local Similarity 29.3%; Pred. No. 1.8e-11;
Matches 49; Conservative 27; Mismatches 52; Indels 39; Gaps 8;

QY 30 TSEVHMIDNGIKQSIGTFTTDTKGLPAGHGFIHEGSGCGPAEHDGHLTAGLQAHGYDPD 103
Db 1 TRAVCVLKGDPVQ--GTTHFEAKGDTVVVVTGSLTGLTEGDHGFHVHQFGDNTQGCTSA 58
QY 83 PRAEHDGHLTAGLQAHGYDPDKTGKHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLT 141
Db 59 P-----HFNP-LSKKGGPKDERHVGDLGNVTADKXGVAIVDIVDPLIS 102
QY 142 VK---EIKGRTVMIH-----AGGDNYSKPLPLGGGGGARIACGVI 178
Db 103 LSGEYSIIIGRTVMVHEKPDGLGRGNEESTK---TGNAGSRLACGVI 146

RESULT 25
US-08-679-493A-205
; Sequence 205, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216

US-09-202-832-16
; Sequence 16, Application US/09202832
; Patent No. 6194190
; GENERAL INFORMATION:
; APPLICANT: IZU, Yukiko
; APPLICANT: TANAKA, Tetsuki
; APPLICANT: MIYAGI, Masaru
; APPLICANT: TANIGAWA, Tetsuo
; APPLICANT: TOMONO, Jun
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: NOVEL AMINO-TERMINAL DEBLOCKING ENZYME
; FILE REFERENCE: 1422-368P
; CURRENT APPLICATION NUMBER: US/09/202,832
; CURRENT FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: superoxide
; OTHER INFORMATION: dismutase
US-09-202-832-16

Query Match      17.5%; Score 166.5; DB 3; Length 151;
Best Local Similarity 29.3%; Pred. No. 1.8e-11;
Matches 49; Conservative 27; Mismatches 52; Indels 39; Gaps 8;

QY 30 TSEVHMIDNGIKQSIGTFTTDTKGLPAGHGFIHEGSGCGPAEHDGHLTAGLQAHGYDPD 103
Db 2 TRAVCVLKGDPVQ--GTTHFEAKGDTVVVVTGSLTGLTEGDHGFHVHQFGDNTQGCTSA 59
QY 83 PRAEHDGHLTAGLQAHGYDPDKTGKHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLT 141
Db 60 P-----HFNP-LSKKGGPKDERHVGDLGNVTADKXGVAIVDIVDPLIS 103
QY 142 VK---EIKGRTVMIH-----AGGDNYSKPLPLGGGGGARIACGVI 178
Db 104 LSGEYSIIIGRTVMVHEKPDGLGRGNEESTK---TGNAGSRLACGVI 147

RESULT 26
US-08-679-493A-205
; Sequence 205, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
```





```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 841 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-709-177-86

Query Match 17.5%; Score 166.5; DB 2; Length 841;
Best Local Similarity 31.9%; Pred. No. 2.2e-10;
Matches 52; Conservative 26; Mismatches 56; Indels 29; Gaps 9;

QY 30 TSEVMIDNGIKSIGVTFTDITDKGLQIKT--DLKGLPAGHGFIHEGSGCGPAEHD 87
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 TNPVCLVKGDPVQ--GIINFEQESNGPVKVGSIKGLTEGLGHGFHVEFGD----- 53
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 88 GHLTAGLQAHG-HYDPDKTGKHEGLG-NGHKGDLPLRVKADGIKAKETLLAPRLTVKE- 144
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 54 --NTAGCTSPGFHNP-LSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVLSLGD 110
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 145 --IKGRTVMIH-----AGDNYSDKPLPLGGGGARIACGVI 178
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 HCIIGRTLIVHEKADDLKGGNEESTK---TGNAGSRLACGVI 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 31
US-08-679-493A-193
; Sequence 193, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 193
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Nicotiana acuminata
; US-08-679-493A-193

Query Match 17.5%; Score 166; DB 3; Length 151;
Best Local Similarity 33.3%; Pred. No. 2e-11;
Matches 48; Conservative 19; Mismatches 55; Indels 22; Gaps 7;

QY 46 GTVTFT-DTDKGLQIKTDLKLPGAGHGFIHEGSGCGPAEHDGHLTAGLQAHG-HYDPD 103
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 GTIIFTQGDAPTTVTGNVSLKPLGLHGFHVALGD-----TTGCMSTGPHYNP- 64
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 104 KTGKHEGLGN--GHKGDPLRVKADGIKAKETLLAPRLTV---KEIKGRTVMIHAGDN 158
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 -AGKEHGAPEDVRHAGDLGNITVGEDGTASTFLTLDKQIPLAGPQSIIGRAVVVHADPDD 123
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 159 YS----DKPLPLGGGGARIACGVI 178
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 LKGGGHELSKTTGNAGRVACGII 147
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 32
US-08-722-050-7
; Sequence 7, Application US/08722050
; Patent No. 5871729
; GENERAL INFORMATION:
```



STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/883,985  
FILING DATE: 20-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIORITY INFORMATION:  
APPLICATION NUMBER: US 09/203,607  
FILING DATE: 02-DEC-1998  
APPLICATION NUMBER: US 08/722,050  
FILING DATE: 23-JAN-1997  
APPLICATION NUMBER: US 08/225,757  
FILING DATE: 11-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: STEFFE, ERIC K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1020003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-883-985-7  
Query Match 17.4%; Score 165.5; DB 4; Length 152;  
Best Local Similarity 31.9%; Pred. No. 2.3e-11;  
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;  
QY 30 TSEVHMIDNGIKQSIGTGTFTDTKGLQIKT--DLKGLPAGEHGHFIHEGSGCGPAEHD 87  
Db 1 TKAVCVLKGDPVQ--GIINFEQKESNGPVKWSIKGLTEGLHGHVHEFGD-----51  
QY 88 GHLTAGLQAHG-HYDPDKTKGHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLTVKE- 144  
Db 52 --NTAGCTSAGPHFN-LSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSD 108  
QY 145 --IKGRTVMIH-----AGGDNYSKPLPLGGGGARIACGVI 178  
Db 109 HCIIGRTLTVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 148  
RESULT 34  
US-08-679-493A-204  
Sequence 204, Application US/08679493A  
Patent No. 6303295  
GENERAL INFORMATION:  
APPLICANT: Taylor, Ethan W.  
TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
FILE REFERENCE: 55-95  
CURRENT FILING DATE: 1996-07-12  
PRIOR FILING DATE: 1995-07-14  
PRIOR APPLICATION NUMBER: 60/001203  
PRIOR FILING DATE: 1995-07-14  
PRIOR APPLICATION NUMBER: 60/003,112  
PRIOR FILING DATE: 1995-09-01  
NUMBER OF SEQ ID NOS: 216  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 204  
LENGTH: 153  
TYPE: PRI  
ORGANISM: Homo sapiens  
US-08-679-493A-204  
Query Match 17.4%; Score 165.5; DB 3; Length 153;  
Best Local Similarity 31.9%; Pred. No. 2.4e-11;  
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;  
QY 30 TSEVHMIDNGIKQSIGTGTFTDTKGLQIKT--DLKGLPAGEHGHFIHEGSGCGPAEHD 87  
Db 2 TKAVCVLKGDPVQ--GIINFEQKESNGPVKWSIKGLTEGLHGHVHEFGD-----52  
QY 88 GHLTAGLQAHG-HYDPDKTKGHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLTVKE- 144  
Db 53 --NTAGCTSAGPHFN-LSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSD 109  
QY 145 --IKGRTVMIH-----AGGDNYSKPLPLGGGGARIACGVI 178  
Db 110 HCIIGRTLTVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 149  
RESULT 35  
US-09-126-109-4  
Sequence 4, Application US/09126109  
Patent No. 6171856  
GENERAL INFORMATION:  
APPLICANT: Thigpen, Anice  
APPLICANT: Hohmeier, Hans-Ewald  
APPLICANT: Newgard, Christopher B.  
APPLICANT: Unger, Roger H.  
APPLICANT: Shimabukuro, Michio  
APPLICANT: Chen, Guaxun  
APPLICANT: Rhodes, Christopher J.  
APPLICANT: Hugl, Sigrun R.  
APPLICANT: Cousin, Sharon  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING  
TO NO-MEDIATED CYTOTOXICITY  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
US-08-679-493A-204  
Query Match 17.4%; Score 165.5; DB 3; Length 153;  
Best Local Similarity 31.9%; Pred. No. 2.4e-11;  
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;  
QY 30 TSEVHMIDNGIKQSIGTGTFTDTKGLQIKT--DLKGLPAGEHGHFIHEGSGCGPAEHD 87  
Db 2 TKAVCVLKGDPVQ--GIINFEQKESNGPVKWSIKGLTEGLHGHVHEFGD-----52  
QY 88 GHLTAGLQAHG-HYDPDKTKGHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLTVKE- 144  
Db 53 --NTAGCTSAGPHFN-LSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSD 109  
QY 145 --IKGRTVMIH-----AGGDNYSKPLPLGGGGARIACGVI 178  
Db 110 HCIIGRTLTVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 149  
RESULT 36  
US-09-126-109-4  
Sequence 4, Application US/09126109  
Patent No. 6171856  
GENERAL INFORMATION:  
APPLICANT: Thigpen, Anice  
APPLICANT: Hohmeier, Hans-Ewald  
APPLICANT: Newgard, Christopher B.  
APPLICANT: Unger, Roger H.  
APPLICANT: Shimabukuro, Michio  
APPLICANT: Chen, Guaxun  
APPLICANT: Rhodes, Christopher J.  
APPLICANT: Hugl, Sigrun R.  
APPLICANT: Cousin, Sharon  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING  
TO NO-MEDIATED CYTOTOXICITY  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas



```
Db 54 --NTAGCTSAGPHNP-LSRKHGPKDEERHVRVGLGNVTADKGDVADVSIEDSVISLSD 110
QY 145 --IKGRTVMIH-----AGGDNYSDKPLPLGGGARIACGVI 178
Db 111 HCLIGRTLTVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 150

RESULT 39
US-07-910-760-12
; Sequence 12, Application US/07910760
; Patent No. 5683864
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/910,760
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1021 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-910-760-12

Query Match 17.4%; Score 165.5; DB 1; Length 1021;
Best Local Similarity 31.9%; Pred. No. 3.8e-10;
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;

QY 30 TSEVHMIDNGIKQSIGTVFTTDTDKGLQIKT--DLKGLPAGHGPHIHEGSGCGPAEHD 87
Db 3 TRAVCVLKGDPVQ--GIINFQKESNGPVGWSIKGLTEGLHGPHVHEFGD-----53
QY 88 GHLTAGLQAHG-HYDPDKTGKHEGLG-NGHKDGLPRLVVKADGIKAKETLLAPRLTVKE-144
Db 54 --NTAGCTSAGPHNP-LSRKHGPKDEERHVRVGLGNVTADKGDVADVSIEDSVISLSD 110
QY 145 --IKGRTVMIH-----AGGDNYSDKPLPLGGGARIACGVI 178
Db 111 HCLIGRTLTVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 150

RESULT 40
US-08-440-519-12
; Sequence 12, Application US/08440519
; Patent No. 5712087
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
```

```
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,519
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1021 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-519-12

Query Match 17.4%; Score 165.5; DB 1; Length 1021;
Best Local Similarity 31.9%; Pred. No. 3.8e-10;
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;

QY 30 TSEVHMIDNGIKQSIGTVFTTDTDKGLQIKT--DLKGLPAGHGPHIHEGSGCGPAEHD 87
Db 3 TRAVCVLKGDPVQ--GIINFQKESNGPVGWSIKGLTEGLHGPHVHEFGD-----53
QY 88 GHLTAGLQAHG-HYDPDKTGKHEGLG-NGHKDGLPRLVVKADGIKAKETLLAPRLTVKE-144
Db 54 --NTAGCTSAGPHNP-LSRKHGPKDEERHVRVGLGNVTADKGDVADVSIEDSVISLSD 110
QY 145 --IKGRTVMIH-----AGGDNYSDKPLPLGGGARIACGVI 178
Db 111 HCLIGRTLTVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 150

Search completed: October 26, 2004, 09:46:23
Job time : 42 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 09:45:05 ; Search time 126 seconds  
(without alignments)  
462.512 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 950

Sequence: 1 MKIKLFFVTSIVTISLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 385   | 40.5        | 186    | 14 | US-10-320-800-60     |
| 2          | 180   | 18.9        | 211    | 16 | US-10-437-963-111088 |
| 3          | 179   | 18.8        | 166    | 15 | US-10-424-599-268560 |
| 4          | 176.5 | 18.6        | 150    | 9  | US-09-883-985-11     |
| 5          | 175.5 | 18.5        | 244    | 16 | US-10-767-701-42073  |
| 6          | 172.5 | 18.2        | 151    | 9  | US-09-883-985-5      |
| 7          | 172.5 | 18.2        | 158    | 15 | US-10-425-114-46108  |
| 8          | 170   | 17.9        | 204    | 15 | US-10-424-599-235729 |
| 9          | 168   | 17.7        | 153    | 9  | US-09-883-985-3      |
| 10         | 167.5 | 17.6        | 154    | 16 | US-10-408-765A-185   |
| 11         | 166.5 | 17.5        | 150    | 9  | US-09-883-985-4      |
| 12         | 166.5 | 17.5        | 151    | 9  | US-09-883-985-6      |
| 13         | 166.5 | 17.5        | 153    | 15 | US-10-425-114-56350  |

|    |       |      |      |    |                      |                    |
|----|-------|------|------|----|----------------------|--------------------|
| 14 | 166.5 | 17.5 | 841  | 10 | US-09-884-456-86     | Sequence 86, Appl  |
| 15 | 166.5 | 17.5 | 841  | 10 | US-09-884-456-86     | Sequence 86, Appl  |
| 16 | 165.5 | 17.4 | 152  | 9  | US-09-883-985-7      | Sequence 7, Appl   |
| 17 | 165.5 | 17.4 | 153  | 17 | US-10-700-816-17     | Sequence 17, Appl  |
| 18 | 165.5 | 17.4 | 154  | 9  | US-09-904-987-6      | Sequence 6, Appl   |
| 19 | 165.5 | 17.4 | 236  | 14 | US-10-272-459-46     | Sequence 46, Appl  |
| 20 | 165.5 | 17.4 | 352  | 14 | US-10-272-459-45     | Sequence 45, Appl  |
| 21 | 165.5 | 17.4 | 382  | 14 | US-10-272-459-47     | Sequence 47, Appl  |
| 22 | 165.5 | 17.4 | 652  | 14 | US-10-272-459-48     | Sequence 48, Appl  |
| 23 | 165.5 | 17.4 | 1099 | 9  | US-09-881-654-4      | Sequence 4, Appl   |
| 24 | 165.5 | 17.4 | 1099 | 15 | US-10-637-323-4      | Sequence 4, Appl   |
| 25 | 165.5 | 17.4 | 1099 | 16 | US-10-658-782-6      | Sequence 6, Appl   |
| 26 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-48136  | Sequence 48136, A  |
| 27 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-52073  | Sequence 52073, A  |
| 28 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-52143  | Sequence 52143, A  |
| 29 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-59106  | Sequence 59106, A  |
| 30 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-61368  | Sequence 61368, A  |
| 31 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-62898  | Sequence 62898, A  |
| 32 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-66160  | Sequence 66160, A  |
| 33 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-67821  | Sequence 67821, A  |
| 34 | 164.5 | 17.3 | 153  | 15 | US-10-425-114-72460  | Sequence 72460, A  |
| 35 | 163.5 | 17.2 | 152  | 16 | US-10-437-963-105229 | Sequence 105229, A |
| 36 | 163.5 | 17.2 | 152  | 16 | US-10-767-701-46195  | Sequence 46195, A  |
| 37 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-44873  | Sequence 44873, A  |
| 38 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-44881  | Sequence 44881, A  |
| 39 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-47023  | Sequence 47023, A  |
| 40 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-47474  | Sequence 47474, A  |
| 41 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-48227  | Sequence 48227, A  |
| 42 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-48230  | Sequence 48230, A  |
| 43 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-58338  | Sequence 58338, A  |
| 44 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-58976  | Sequence 58976, A  |
| 45 | 163.5 | 17.2 | 153  | 15 | US-10-425-114-60167  | Sequence 60167, A  |

ALIGNMENTS

RESULT 1  
US-10-320-800-60  
; Sequence 60, Application US/10320800  
; Publication No. US20030215469A1  
; GENERAL INFORMATION:  
; APPLICANT: ROBINSON, ANDREW  
; APPLICANT: GORRINGE, ANDREW  
; APPLICANT: HUDSON, MICHAEL  
; APPLICANT: REDDIN, KAREN  
; TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE  
; FILE REFERENCE: 1581.0790001  
; CURRENT APPLICATION NUMBER: US/10/320,800  
; CURRENT FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: PCT/GB99/03626  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 60  
; LENGTH: 186  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-10-320-800-60

|                       |       |  |                |             |
|-----------------------|-------|--|----------------|-------------|
| Query Match           | 40.5% | Score 385;   | DB 14;         | Length 186; |
| Best Local Similarity | 50.3% | Pred. No. 3.4e-32;   |                |             |
| Matches               | 77;   | Conservative 22;   | Mismatches 52; | Gaps 2;     |
| QY                    | 28    | SVTSEVHMDDNGIKQSGTGTFTTDTDKGLQKTDKGLPAGHGFHIEGSGCGPAEHD      | 87             |             |
| DB                    | 33    | STEVKVVQQLDPYNGNKDVGTVTTESNYGLVPTDQLGLSEGLHGFHIEHNSCFEKE     | 92             |             |
| QY                    | 88    | GHLTLAGLQAHGHYDPPDKTKGEGP-LGNCHKGLDRLVVKADGIKATKLLAPRLT-VKEI | 145            |             |
| DB                    | 93    | GKLTLAGLQAHGHYDPPDKTKGEGP-LGNCHKGLDRLVVKADGIKATKLLAPRLT-VKEI | 152            |             |
| QY                    | 146   | KGRVTMIHAGSDNYSDKPLPLPLGGGGARIACGVI                          | 178            |             |



```
QY 101 DPDKTKGHEGPIGNHGKSDPLRLVVKADGIKAKETLLAPRLTVK---EIKGRTVMIH-----153
Db 71 SPKADRHVGDGLN-----VTAEGGVAQFNFTDQISLKGRSIIIGRTAVVHEKQD 121

QY 154 ----AGDNYSDKPLPLGGGARIACGVI 178
Db 122 DLGKGGD---DESLKTNAGGRLACGVI 146

RESULT 5
US-10-767-701-42073
; Sequence 42073, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42073
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C17960_1.pep
US-10-767-701-42073

Query Match 18.5%; Score 175.5; DB 16; Length 244;
Best Local Similarity 30.5%; Pred. No. 5.3e-10;
Matches 51; Conservative 21; Mismatches 48; Indels 47; Gaps 8;

QY 30 TSEVMIDDNGIKQSIGTVFTTDTKG-LQIKTDLKLPGAGHGHG-----SC 81
Db 102 TSEVE-----GVVTLTQDDGPTTVNVRITGLTPGLHGFHLHFGDTTNGCIST 150

QY 82 GPAEHDGHLTAGLQAHGHYDPDKTGHEGPIGNHGKSDPLRLVVKADGIKAKETLLAPRLT 141
Db 151 GPHFNPNLIT-----HGAPDEVHRHAGDLN-----IVANAEGVAEATIVDTQIP 195

QY 142 V---KEIKGRTVMIH-----AGDNYSDKPLPLGGGARIACGVI 178
Db 196 LSGPNSVVGRAFWVHELEDDLKGKGHELS---LSTGNAGRLACGVV 239

RESULT 6
US-09-883-985-5
; Sequence 5, Application US/09883985
; Patent No. US20020081288A1
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: FRASER, CLAIRE M.
; APPLICANT: GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-883-985-5

Query Match 18.2%; Score 172.5; DB 9; Length 151;
Best Local Similarity 33.6%; Pred. No. 5.9e-10;
Matches 50; Conservative 22; Mismatches 54; Indels 23; Gaps 9;

QY 40 GIKSIGTVFTTDTKGLQIKT-DLKGLPAGHGHGPHIHEGSGCPAEHDGHLTAGLQAHG 98
Db 12 GVK---GTIFFTHEGNGATTVGTGSLRPLGLHGFHVHALGD---NTNGCMSTG---P 60

QY 99 HYDPDKTKG-HEGPI-GNGHKGDPLRLVVKADGIKAKETLLAPRLTV---KEIKGRTVMIH 153
Db 61 HFNPDP--GKTHGAPEDANRHAGDLGNIIVGDDGTATFTTDSQIPLSGPNSIVGRAIVVH 118

QY 154 AGGDNYV---DKPLPLGGGARIACGVI 178
Db 119 ADPDDLKGKGHELSLSTGNAGRVACGII 147

RESULT 7
US-10-425-114-46108
; Sequence 46108, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46108
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701177211_FLI.pgp
US-10-425-114-46108

Query Match 18.2%; Score 172.5; DB 15; Length 158;
Best Local Similarity 31.3%; Pred. No. 6.2e-10;
Matches 46; Conservative 23; Mismatches 49; Indels 29; Gaps 8;
```





QY 46 GTVTFTDDKGLQIKT--DLKGLPAGEHGFIHEGSGCGPAEHDCGHLTAGIQAAGHVDPD 103

```
Db 13 GTVFFQESSGTPVKVSEVCLAKLGHGFVHFHFGD-----NTNGCMSSG-----PHFNP- 63
Qy 104 KTGK-HEGPLG-NGHKGLDPLRVKADGIAKETLLAPRLTV---KEIKGRVTMIHAGGDN 158
Db 64 -YGKEGAPVDENRHLGLDGNIEATGDCPTKVNITDSKITLFGADSIIGRTVVVHADADD 122
Qy 159 YSDKPLPL-----GGGGARIAAGVVI 178
Db 123 LGQGGHLSKSTGNAGARIGCGVI 146

RESULT 13
US-10-425-114-56350
; Sequence 56350, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56350
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73194H01_FLI.pep
US-10-425-114-56350

Query Match 17.5%; Score 166.5; DB 15; Length 153;
Best Local Similarity 32.0%; Pred. No. 2.5e-09;
Matches 49; Conservative 20; Mismatches 55; Indels 29; Gaps 9;

Qy 39 NGIKOSIGTVFTDIDKG-LQIKTDLKGLPAGEHGFHHEGSGCPAEHDGHLTAGLOAH 97
Db 13 DGVR---GTIFTOEGDGTTVTGSVGLKPLGLHGFVHAGD-----TTNGCMST 60

Qy 98 G-HYDDPKTGKHEGPLG-NGHKGLDPLRVKADGIAKETLLAPRLTV---KEIKGRVTMI 152
Db 61 GPHYNP-ASKHEGAPEDENRHAGDLGNVTAGDGVANINVTDTSQIPLTGPNISIGRAVV 119

Qy 153 HA-----GGDNYSDKPLPLGGGGARIAAGVVI 178
Db 120 HADPDLDLKGGHFLKSP---GNAGGRVACGII 149
```

```
RESULT 14
US-09-884-456-86
; Sequence 86, Application US/09884456
; Publication No. US20030027317A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Oui-Lim
; TITLE OF INVENTION: Hepatitis C virus protease
; FILE REFERENCE: 22302010005
; CURRENT APPLICATION NUMBER: US/09/884,456
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/253,230
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: 08/709,177
; PRIOR FILING DATE: 1996-09-06
; PRIOR APPLICATION NUMBER: 08/709,177
; PRIOR FILING DATE: 1995-05-12
; PRIOR APPLICATION NUMBER: 08/350,884
; PRIOR FILING DATE: 1994-12-06
```

```
; PRIOR APPLICATION NUMBER: 07/680,296
; PRIOR FILING DATE: 1991-04-04
; PRIOR APPLICATION NUMBER: 07/505,433
; PRIOR FILING DATE: 1990-04-04
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vector cflSODp600
US-09-884-456-86

Query Match 17.5%; Score 166.5; DB 10; Length 841;
Best Local Similarity 31.9%; Pred. No. 2.4e-08;
Matches 52; Conservative 26; Mismatches 56; Indels 29; Gaps 9;

Qy 30 TSEVHMIDDNGIKOSIGTVFTDIDKGLOIKT--DLKGLPAGEHGFHHEGSGCPAEHD 87
Db 3 TNPVCLKGDGPVQ--GIINFEQESNGPVKVGSIKGLTEGLHGFVHFEGD-----53

Qy 88 GHILTAGLOAHG-HYDDPKTGKHEGPLG-NGHKGLDPLRVKADGIAKETLLAPRLTVKE- 144
Db 54 --NTAGCTSPGFHNP-LSRKHGQPKDEERHVGDLGNVTADKDGADVDSIEDSVISLGD 110

Qy 145 --IKGRVTMIH-----AGDNYSDKPLPLGGGGARIAAGVVI 178
Db 111 HCLIGRLIVHEKADDLGKGNEESTK---TGNAGSRLACGVI 150

RESULT 15
US-09-884-455-86
; Sequence 86, Application US/09884455
; Publication No. US20030064499A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Oui-Lim
; TITLE OF INVENTION: Hepatitis C virus protease
; FILE REFERENCE: 22302010004
; CURRENT APPLICATION NUMBER: US/09/884,455
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/253,675
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 08/709,177
; PRIOR FILING DATE: 1996-09-06
; PRIOR APPLICATION NUMBER: 08/440,548
; PRIOR FILING DATE: 1995-05-12
; PRIOR APPLICATION NUMBER: 08/350,884
; PRIOR FILING DATE: 1994-12-06
; PRIOR APPLICATION NUMBER: 07/680,296
; PRIOR FILING DATE: 1991-04-04
; PRIOR APPLICATION NUMBER: 07/505,433
; PRIOR FILING DATE: 1990-04-04
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vector cflSODp600
US-09-884-455-86

Query Match 17.5%; Score 166.5; DB 10; Length 841;
Best Local Similarity 31.9%; Pred. No. 2.4e-08;
Matches 52; Conservative 26; Mismatches 56; Indels 29; Gaps 9;

Qy 30 TSEVHMIDDNGIKOSIGTVFTDIDKGLOIKT--DLKGLPAGEHGFHHEGSGCPAEHD 87
Db 3 TNPVCLKGDGPVQ--GIINFEQESNGPVKVGSIKGLTEGLHGFVHFEGD-----53
```

|    |     |  |     |
|----|-----|--|-----|
| Qy | 88  | GHLTACLOAHG--HYDPDGTGHEGPLG--NGHKGDLPRLVVKADGIAKETLAPLTVKE-- | 144 |
| Db | 54  | --NTAGCTSPGFHNP--LSRKHGGPKDEERHVGDLGNVTADKGVADVSDSVLSLGD     | 110 |
| Qy | 145 | --IKRGTVMIH-----AGGDNYSDKPLPLGGGGARIACGVI                    | 178 |
| Db | 111 | HGCIIGRTIIVVHKADLKGKGNNESTK---TGNAGSRLACGVI                  | 150 |

RESULT 16  
US-09-883-985-7  
; Sequence 7, Application US/09883985  
; Patent No. US2002008128A1  
; GENERAL INFORMATION:  
; APPLICANT: YU, GUO-LIANG  
; ROSEN, CRAIG A.  
; FRASER, CLAIRE M.  
; GOCAYNE, JEANNINE D.  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA

|                       |        |                    |                |                    |
|-----------------------|--------|--------------------|----------------|--------------------|
| Query Match           | 17.4%  | Score 165.5;       | DB 9;          | Length 152;        |
| Best Local Similarity | 31.9%; | pred. No. 3.2e-09; |                |                    |
| Matches               | 52;    | Conservative 25;   | Mismatches 57; | Indels 29; Gaps 9; |

  

|    |     |   |     |
|----|-----|---|-----|
| QY | 30  | TSEVHMDDNGIKSIGITVTFDTDKGLQIKT--DLKGLPAGEHGFIHEGSGCGPAEHD | 87  |
|    |     | :           :           :                                 |     |
| Db | 1   | TKAVCVLKGDPVQ--GIINFQEKESGPNVWGSIKGLTEGLAGFHVHEFGD-----   | 51  |
|    |     | :           :           :                                 |     |
| QY | 88  | GHLTAGLQAHG-HYDDPKTKGHEGLG-NGHKGLPLRVVKADGIAKETLLAPLTVKE- | 144 |
|    |     | :           :           :                                 |     |
| Db | 52  | --NTAGCTAGPHNP-LSRKHGPKDEHRHVGLGNVTADKQGVADVIEDSVISLSD    | 108 |
|    |     | :           :           :                                 |     |
| QY | 145 | --IKGRVTMIH-----AGDNDYSDKPLPLGGGGGARIACGVI                | 178 |
|    |     | :           :           :                                 |     |

```

Db      109 HCIIGRTLTVVHEKADDLGKGNEESTK---TGNAGSRILACGVI 148
      |||:::| ||:| | | | | | | | | | | | | | | | | | | | |
RESULT 17
US-10-700-816-17
; Sequence 17, Application US/10700816
; Publication No. US20040192629A1

```

|                       |        |                         |                                       |                    |
|-----------------------|--------|-------------------------|---------------------------------------|--------------------|
| Query Match           | 17.4%; | Score 165.5;            | DB 17;                                | Length 153;        |
| Best Local Similarity | 31.9%; | Pred. No. 3.2e-09;      |                                       |                    |
| Matches               | 52;    | Conservative 25;        | Mismatches 57;                        | Indels 29; Gaps 9; |
| QY                    | 30     | TSEVHMIDNGIKSIGTGTFTD   | DKGLQIKT--DLKGLPAGRHGHPIHGGSCGPAEHD   | 87                 |
| Db                    | 2      | TKAVCVLKGDPVQ--GI       | INFQESGNPVKWSIKGLTEGHLGHVHFEGD        | 52                 |
| QY                    | 88     | GHLTGLQAHG-HYDPDKTCKHEG | PLG-NGHKGDLPRLVVKADGIAKETILLAPRLTVKE- | 144                |
| Db                    | 53     | --NTACTSAGPHFNP-LSRKGG  | PKDEERHVGDLGNVTADKGVADVSTEDSVISLGD    | 109                |
| QY                    | 145    | --IKGTVMTH-----         | AGDNYSKPLPLGGGARIACGVI                | 178                |
| Db                    | 110    | HCIIIGRTLTVVHKADDLGKGN  | PEESTK---TGNAGSRLACGVI                | 149                |

```

RESULT 18
US-09-904-987-6
; Sequence 6, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908A1actyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Con
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 154
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / P00441
; DATABASE ENTRY DATE: 2000-05-30
; RELEVANT RESIDUES: (1) .. (154)
US-09-904-987-6

```

```

Query Match      17.4%; Score 165.5; DB 9; Length 154;
Best Local Similarity 31.9%; Pred. No. 3.3e-09;
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;

QY 30 TSEVHMIDNNGIKQSIGTWTFTDTRGLQIKT--DLKGLPAGHGPHIIEGSGCGPAEHD 87
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3  TRAVCVLKGDPGYO--GIINFEEKSNPGPVKWSIGKLTGELHGPHVFEFGD-----53

```



; OTHER INFORMATION: of human superoxide dismutase fused with the HAV  
; OTHER INFORMATION: nonstructural protein  
US-10-272-459-48

Query Match 17.4%; Score 165.5; DB 14; Length 652;  
Best Local Similarity 31.9%; Pred. No. 2.2e-08;  
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;  
QY 30 TSEVMIDNGIKQSIGTFTDTDKGLQIKT--DLKGLPAGEHGFHIEGSGCGPAEHD 87  
Db 3 TKAVCVLKGDPVQ--GIINFQKESNGPVKWSIKGLTEGLHGFHVEFGD-----53  
QY 88 GHLTAGLQAHG-HYDPDKTKHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLTVKE- 144  
Db 54 --NTAGCTSAGPHFNP-LSRKHGGPKDERHVGDLGNVTADKGDVADVSIEDSVISLSD 110  
QY 145 --IKGRTVMIH-----AGGDNYSKPLPLGGGGARIACGVI 178  
Db 111 HCLIGRTLVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 150

RESULT 23  
US-09-881-654-4  
; Sequence 4, Application US/09891654  
; Patent No. US20020146685A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIEN, David Y.  
; APPLICANT: ARCANDEL, Phillip  
; APPLICANT: TANDESKE, Laura  
; APPLICANT: GEORGE-NASCIEMENTO, Carlos  
; APPLICANT: COLT, Doris  
; APPLICANT: MEDINA-SELBY, Angelica  
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES  
; FILE REFERENCE: 2302-17039 / PP17039.002  
; CURRENT APPLICATION NUMBER: US/09/881,654  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/212,082  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/280,811  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/280,867  
; PRIOR FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1099  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: MEPA 7.1  
US-09-881-654-4

Query Match 17.4%; Score 165.5; DB 9; Length 1099;  
Best Local Similarity 31.9%; Pred. No. 4.4e-08;  
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;  
QY 30 TSEVMIDNGIKQSIGTFTDTDKGLQIKT--DLKGLPAGEHGFHIEGSGCGPAEHD 87  
Db 3 TKAVCVLKGDPVQ--GIINFQKESNGPVKWSIKGLTEGLHGFHVEFGD-----53  
QY 88 GHLTAGLQAHG-HYDPDKTKHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLTVKE- 144  
Db 54 --NTAGCTSAGPHFNP-LSRKHGGPKDERHVGDLGNVTADKGDVADVSIEDSVISLSD 110  
QY 145 --IKGRTVMIH-----AGGDNYSKPLPLGGGGARIACGVI 178  
Db 111 HCLIGRTLVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 150

RESULT 24  
US-10-637-323-4  
; Sequence 4, Application US/10637323  
; Publication No. US20040063092A1

; GENERAL INFORMATION:  
; APPLICANT: CHIEN, David Y.  
; APPLICANT: ARCANDEL, Phillip  
; APPLICANT: TANDESKE, Laura  
; APPLICANT: GEORGE-NASCIEMENTO, Carlos  
; APPLICANT: COLT, Doris  
; APPLICANT: MEDINA-SELBY, Angelica  
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES  
; FILE REFERENCE: 2302-17039 / PP17039.002  
; CURRENT APPLICATION NUMBER: US/10/637,323  
; CURRENT FILING DATE: 2003-08-08  
; PRIOR APPLICATION NUMBER: US/09/881,654  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/212,082  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/280,811  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/280,867  
; PRIOR FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1099  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: MEPA 7.1  
US-10-637-323-4

Query Match 17.4%; Score 165.5; DB 15; Length 1099;  
Best Local Similarity 31.9%; Pred. No. 4.4e-08;  
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;  
QY 30 TSEVMIDNGIKQSIGTFTDTDKGLQIKT--DLKGLPAGEHGFHIEGSGCGPAEHD 87  
Db 3 TKAVCVLKGDPVQ--GIINFQKESNGPVKWSIKGLTEGLHGFHVEFGD-----53  
QY 88 GHLTAGLQAHG-HYDPDKTKHEGPLG-NGHKGDLPRLVVKADGIKAKETLLAPRLTVKE- 144  
Db 54 --NTAGCTSAGPHFNP-LSRKHGGPKDERHVGDLGNVTADKGDVADVSIEDSVISLSD 110  
QY 145 --IKGRTVMIH-----AGGDNYSKPLPLGGGGARIACGVI 178  
Db 111 HCLIGRTLVVHEKADDLGKGGNEESTK---TGNAGSRLACGVI 150

RESULT 25  
US-10-658-782-6  
; Sequence 6, Application US/10658782  
; Publication No. US20040142321A1  
; GENERAL INFORMATION:  
; APPLICANT: ARCANDEL, Phillip  
; APPLICANT: CHIEN, David Y.  
; TITLE OF INVENTION: HCV ASSAY  
; FILE REFERENCE: 2300-19199  
; CURRENT APPLICATION NUMBER: US/10/658,782  
; CURRENT FILING DATE: 2003-09-08  
; PRIOR APPLICATION NUMBER: 60/409,515  
; PRIOR FILING DATE: 2002-09-09  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In version 3.2  
; SEQ ID NO 6  
; LENGTH: 1099  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: MEPA 7.1 amino acid sequence  
US-10-658-782-6

Query Match 17.4%; Score 165.5; DB 16; Length 1099;  
Best Local Similarity 31.9%; Pred. No. 4.4e-08;  
Matches 52; Conservative 25; Mismatches 57; Indels 29; Gaps 9;



```
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 59106
/ LENGTH: 153
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4074-013-H7_FLI.pep
US-10-425-114-59106

Query Match      17.3%; Score 164.5; DB 15; Length 153;
Best Local Similarity 32.0%; Pred. No. 4.1e-09;
Matches 49; Conservative 20; Mismatches 55; Indels 29; Gaps 9;

QY 39 NGIKQSIGVTFTDQK-LQIKDLKGLPAGEHGFHIEGSCGPAEDHGLTAGLQAH 97
Db 13 DGVK--GTFITQVGDGPTTVTGSVGLKPLGPHFVHALGD-----TTNGCMST 60

QY 98 G-HYDPDKTGKEGPLG-NGHKDGLPRLVVKADGIKETLLAPRLTV---KEIKGRTYMI 152
Db 61 GPHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANINVTDSQIPLTGPNSIIGRAVVV 119

QY 153 HA-----GGDNYSKPLPLGGGGARIACGVI 178
Db 120 HADPDDLKGKGHELSKS---TGNAGGRVACGII 149

RESULT 30
US-10-425-114-61368
/ Sequence 61368, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 61368
/ LENGTH: 153
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3061-037-H2_FLI.pep
US-10-425-114-61368

Query Match      17.3%; Score 164.5; DB 15; Length 153;
Best Local Similarity 32.0%; Pred. No. 4.1e-09;
Matches 49; Conservative 20; Mismatches 55; Indels 29; Gaps 9;

QY 39 NGIKQSIGVTFTDQK-LQIKDLKGLPAGEHGFHIEGSCGPAEDHGLTAGLQAH 97
Db 13 DGVK--GTFITQVGDGPTTVTGSVGLKPLGPHFVHALGD-----TTNGCMST 60

QY 98 G-HYDPDKTGKEGPLG-NGHKDGLPRLVVKADGIKETLLAPRLTV---KEIKGRTYMI 152
Db 61 GPHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANINVTDSQIPLTGPNSIIGRAVVV 119

QY 153 HA-----GGDNYSKPLPLGGGGARIACGVI 178
Db 120 HADPDDLKGKGHELSKS---TGNAGGRVACGII 149
```

```
Db 120 HADPDDLKGKGHELSKS---TGNAGGRVACGII 149

RESULT 31
US-10-425-114-62898
/ Sequence 62898, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 62898
/ LENGTH: 153
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB143-041-B12_FLI.pep
US-10-425-114-62898

Query Match      17.3%; Score 164.5; DB 15; Length 153;
Best Local Similarity 32.0%; Pred. No. 4.1e-09;
Matches 49; Conservative 20; Mismatches 55; Indels 29; Gaps 9;

QY 39 NGIKQSIGVTFTDQK-LQIKDLKGLPAGEHGFHIEGSCGPAEDHGLTAGLQAH 97
Db 13 DGVK--GTFITQVGDGPTTVTGSVGLKPLGPHFVHALGD-----TTNGCMST 60

QY 98 G-HYDPDKTGKEGPLG-NGHKDGLPRLVVKADGIKETLLAPRLTV---KEIKGRTYMI 152
Db 61 GPHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANINVTDSQIPLTGPNSIIGRAVVV 119

QY 153 HA-----GGDNYSKPLPLGGGGARIACGVI 178
Db 120 HADPDDLKGKGHELSKS---TGNAGGRVACGII 149

RESULT 32
US-10-425-114-66160
/ Sequence 66160, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 66160
/ LENGTH: 153
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4605-009-C9_FLI.pep
US-10-425-114-66160

Query Match      17.3%; Score 164.5; DB 15; Length 153;
Best Local Similarity 32.0%; Pred. No. 4.1e-09;
```

| Matches  | 49; | Conservative  | 20; | Mismatches | 55; | Indels | 29; | Gaps | 9; |
|--|-----|---|-----|------------|-----|--------|-----|------|----|
| Qy   | 39  | NGIKQSIGTVFTTDDKG-LQIKTDLKGLPAGEHGFIHGGSCGPAEHDGHLTAGLOAH     | 97  | :          | :   | :      | :   | :    | :  |
| Db   | 13  | DGVK---GTIFFTQVGDDPTTGVSGSLKPLGHGFVHALGD-----TTNGCMST         | 60  | :          | :   | :      | :   | :    | :  |
| Qy   | 98  | G-HYDPDKTGKHEGPLG-NGHKGDLPLRVVKGADGIKAKETILAPLTV---KEIKORTYMI | 152 | :          | :   | :      | :   | :    | :  |
| Db   | 61  | GPHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANINVTDSQIPLTGPNSIIGRAVVV  | 119 | :          | :   | :      | :   | :    | :  |
| Qy   | 153 | HA-----GGDNYSDKPLPLGGGGARIACGVI                               | 178 | :          | :   | :      | :   | :    | :  |
| Db   | 120 | HADPDDLKGKGHELSKS---TGNAGGRVACGII                             | 149 | :          | :   | :      | :   | :    | :  |
| RESULT 33  |     |   |     |            |     |        |     |      |    |
| US-10-425-114-67821  |     |   |     |            |     |        |     |      |    |
| ; Sequence 67821, Application US/10425114  |     |   |     |            |     |        |     |      |    |
| ; Publication No. US20040034888A1  |     |   |     |            |     |        |     |      |    |
| ; GENERAL INFORMATION:   |     |   |     |            |     |        |     |      |    |
| ; APPLICANT: Liu, Jingdong   |     |   |     |            |     |        |     |      |    |
| ; APPLICANT: Zhou, Yihua   |     |   |     |            |     |        |     |      |    |
| ; APPLICANT: Kovalic, David K.   |     |   |     |            |     |        |     |      |    |
| ; APPLICANT: Screen, Steven E  |     |   |     |            |     |        |     |      |    |
| ; APPLICANT: Tabaska, Jack E   |     |   |     |            |     |        |     |      |    |
| ; APPLICANT: Cao, Yongwei  |     |   |     |            |     |        |     |      |    |
| ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With |     |   |     |            |     |        |     |      |    |
| ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement              |     |   |     |            |     |        |     |      |    |
| ; FILE REFERENCE: 38-21(53313)B  |     |   |     |            |     |        |     |      |    |
| ; CURRENT APPLICATION NUMBER: US/10/425,114                                      |     |   |     |            |     |        |     |      |    |
| ; CURRENT FILING DATE: 2003-04-28  |     |   |     |            |     |        |     |      |    |
| ; NUMBER OF SEQ ID NOS: 73128  |     |   |     |            |     |        |     |      |    |
| ; SEQ ID NO 67821  |     |   |     |            |     |        |     |      |    |
| ; LENGTH: 153  |     |   |     |            |     |        |     |      |    |
| ; TYPE: PRT  |     |   |     |            |     |        |     |      |    |
| ; ORGANISM: Zea mays   |     |   |     |            |     |        |     |      |    |
| ; FEATURE:   |     |   |     |            |     |        |     |      |    |
| ; OTHER INFORMATION: Clone ID: LIB3732-001-H3_FLI.pep                            |     |   |     |            |     |        |     |      |    |
| US-10-425-114-67821  |     |   |     |            |     |        |     |      |    |
| Query Match 17.3%; Score 164.5; DB 15; Length 153;                               |     |   |     |            |     |        |     |      |    |
| Best Local Similarity 32.0%; Pred. No. 4.1e-09;                                  |     |   |     |            |     |        |     |      |    |
| Matches  | 49; | Conservative  | 20; | Mismatches | 55; | Indels | 29; | Gaps | 9; |
| Qy   | 39  | NGIKQSIGTVFTTDDKG-LQIKTDLKGLPAGEHGFIHGGSCGPAEHDGHLTAGLOAH     | 97  | :          | :   | :      | :   | :    | :  |
| Db   | 13  | DGVK---GTIFFTQVGDDPTTGVSGSLKPLGHGFVHALGD-----TTNGCMST         | 60  | :          | :   | :      | :   | :    | :  |
| Qy   | 98  | G-HYDPDKTGKHEGPLG-NGHKGDLPLRVVKGADGIKAKETILAPLTV---KEIKORTYMI | 152 | :          | :   | :      | :   | :    | :  |
| Db   | 61  | GPHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANINVTDSQIPLTGPNSIIGRAVVV  | 119 | :          | :   | :      | :   | :    | :  |
| Qy   | 153 | HA-----GGDNYSDKPLPLGGGGARIACGVI                               | 178 | :          | :   | :      | :   | :    | :  |
| Db   | 120 | HADPDDLKGKGHELSKS---TGNAGGRVACGII                             | 149 | :          | :   | :      | :   | :    | :  |
| RESULT 34  |     |   |     |            |     |        |     |      |    |
| US-10-425-114-72460  |     |   |     |            |     |        |     |      |    |



```
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46195
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2913_1.pep
US-10-767-701-46195

Query Match      17.2%; Score 163.5; DB 16; Length 152;
Best Local Similarity 32.2%; Pred. No. 5.2e-09;
Matches 49; Conservative 20; Mismatches 54; Indels 29; Gaps 9;

QY 40 GIKQSIGTVTTDDTKG-LQIKTDLKGHPAGHGFHIEGGSCGPAEHGHLTAGLQAHG 98
Db 13 GVK---GTIFFQESGDPPTVTVGSVSLKPLGPHGFHVALGD-----TTNGCMSTG 60

QY 99 -HYDPDKTGKHEGPLG-NGHKGDLPRLVVKADGIAKETLLAPRLTV---KEIKGRTVMIH 153
Db 61 PHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANISVTDQSQPLTGPNISIGRAVVVH 119

QY 154 A-----GGDNYSKPLPLGGGGARIACGVI 178
Db 120 ADPDDLKGKGHELSKS---TGNAGGRVACGII 148

RESULT 37
US-10-425-114-44873
; Sequence 44873, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44873
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700618619_FLI.pep
US-10-425-114-44873

Query Match      17.2%; Score 163.5; DB 15; Length 153;
Best Local Similarity 32.0%; Pred. No. 5.3e-09;
Matches 49; Conservative 20; Mismatches 55; Indels 29; Gaps 9;

QY 39 NGIKSIGTVTTDDTKG-LQIKTDLKGHPAGHGFHIEGGSCGPAEHGHLTAGLQAH 97
Db 13 DGVK---GTIFFQESGDPPTVTVGSVSLKPLGPHGFHVALGD-----TTNGCMST 60

QY 98 G-HYDPDKTGKHEGPLG-NGHKGDLPRLVVKADGIAKETLLAPRLTV---KEIKGRTYMI 152
Db 61 GHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANINVTDSQIPLTGPNISIGRAVVV 119
```

```
QY 153 HA-----GGDNYSKPLPLGGGGARIACGVI 178
Db 120 HADPDDLKGKGHELSKS---TGNAGGRVACGII 149

RESULT 38
US-10-425-114-44881
; Sequence 44881, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44881
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700580977_FLI.pep
US-10-425-114-44881

Query Match      17.2%; Score 163.5; DB 15; Length 153;
Best Local Similarity 32.0%; Pred. No. 5.3e-09;
Matches 49; Conservative 20; Mismatches 55; Indels 29; Gaps 9;

QY 39 NGIKSIGTVTTDDTKG-LQIKTDLKGHPAGHGFHIEGGSCGPAEHGHLTAGLQAH 97
Db 13 DGVK---GTIFFQESGDPPTVTVGSVSLKPLGPHGFHVALGD-----TTNGCMST 60

QY 98 G-HYDPDKTGKHEGPLG-NGHKGDLPRLVVKADGIAKETLLAPRLTV---KEIKGRTYMI 152
Db 61 GHYNP-ASKEHGAPEDENRHAGDLGNVTAGADGVANINVTDSQIPLTGPNISIGRAVVV 119

QY 153 HA-----GGDNYSKPLPLGGGGARIACGVI 178
Db 120 HADPDDLKGKGHELSKS---TGNAGGRVACGII 149

RESULT 39
US-10-425-114-47023
; Sequence 47023, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47023
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700344278_FLI.pep
US-10-425-114-47023

Query Match      17.2%; Score 163.5; DB 15; Length 153;
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 09:45:45 ; Search time 154 Seconds  
(without alignments)  
419.294 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 180

Sequence: 1 MKIKLFFVTSIVTISLLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 180   | 100.0       | 180    | 4  | AAB47008 L. intrac |
| 2          | 16    | 8.9         | 213    | 6  | ADA34763 Acinetoba |
| 3          | 9     | 5.0         | 144    | 7  | ABO62994 Klebsiell |
| 4          | 9     | 5.0         | 154    | 2  | AAR32374 20kD Bruc |
| 5          | 8     | 4.4         | 18     | 6  | ABP82477 G protein |
| 6          | 8     | 4.4         | 153    | 5  | ABB77456 Fungi str |
| 7          | 8     | 4.4         | 175    | 7  | ADC00874 Enterohae |
| 8          | 8     | 4.4         | 175    | 7  | ADC00431 Enterohae |
| 9          | 8     | 4.4         | 510    | 6  | ABR57484 Flavobact |
| 10         | 8     | 4.4         | 532    | 1  | AAP96205 Human mus |
| 11         | 8     | 4.4         | 532    | 4  | ABBS6364 Non-endo  |
| 12         | 8     | 4.4         | 532    | 4  | AAU97746 Human CHR |
| 13         | 8     | 4.4         | 532    | 5  | AAU97549 Human cho |
| 14         | 8     | 4.4         | 532    | 6  | ABP81855 Human mus |
| 15         | 8     | 4.4         | 532    | 6  | ABP81855 Human mus |
| 16         | 8     | 4.4         | 532    | 7  | ADD29411 Human mus |
| 17         | 8     | 4.4         | 532    | 7  | ADE40455 Human mus |
| 18         | 8     | 4.4         | 532    | 8  | ADO29257 Human GPC |
| 19         | 8     | 4.4         | 532    | 8  | ADO29258 Mouse GPC |
| 20         | 8     | 4.4         | 533    | 6  | ABR57482 Flavobact |
| 21         | 8     | 4.4         | 684    | 7  | ADC37562 Human nuc |
| 22         | 8     | 4.4         | 759    | 2  | AAW41927 Homo sapi |
| 23         | 8     | 4.4         | 759    | 2  | AAW46593 Human tel |
| 24         | 8     | 4.4         | 759    | 4  | ABO32754 Human pro |
| 25         | 8     | 4.4         | 956    | 6  | ABO53053 Human put |

|    |   |     |     |   |                     |
|----|---|-----|-----|---|---------------------|
| 26 | 7 | 3.9 | 10  | 4 | AAG86501 Saccharom  |
| 27 | 7 | 3.9 | 20  | 5 | ABR73398 IL-1 R an  |
| 28 | 7 | 3.9 | 20  | 7 | ADC99215 Cancer-re  |
| 29 | 7 | 3.9 | 21  | 3 | AAB17944 IL-1 R an  |
| 30 | 7 | 3.9 | 56  | 4 | AAU62481 Propionib  |
| 31 | 7 | 3.9 | 56  | 6 | ABMS5000 Propionib  |
| 32 | 7 | 3.9 | 64  | 4 | AAU44577 Propionib  |
| 33 | 7 | 3.9 | 64  | 6 | ABM41096 Propionib  |
| 34 | 7 | 3.9 | 91  | 2 | AAW27922 Staphyloc  |
| 35 | 7 | 3.9 | 93  | 3 | AAG35331 Zea mays   |
| 36 | 7 | 3.9 | 103 | 4 | AAO11158 Human pol  |
| 37 | 7 | 3.9 | 118 | 5 | ABP08245 Human ORF  |
| 38 | 7 | 3.9 | 123 | 6 | ABU29948 Protein e  |
| 39 | 7 | 3.9 | 124 | 6 | ABM65961 Propionib  |
| 40 | 7 | 3.9 | 124 | 6 | ABU29214 Protein e  |
| 41 | 7 | 3.9 | 125 | 3 | AAG57848 Zea mays   |
| 42 | 7 | 3.9 | 129 | 5 | ABBS0017 Listeria   |
| 43 | 7 | 3.9 | 133 | 7 | ADC96469 E. faeciu  |
| 44 | 7 | 3.9 | 139 | 3 | AAG35329 Zea mays   |
| 45 | 7 | 3.9 | 153 | 3 | ABO72455 Pseudomon  |
| 46 | 7 | 3.9 | 157 | 7 | ADB64754 Human pro  |
| 47 | 7 | 3.9 | 157 | 7 | ADH85826 Enterococ  |
| 48 | 7 | 3.9 | 167 | 3 | AAG57847 Zea mays   |
| 49 | 7 | 3.9 | 178 | 7 | ADF04181 Bacterial  |
| 50 | 7 | 3.9 | 179 | 4 | ABB64175 Drosophil  |
| 51 | 7 | 3.9 | 181 | 4 | ABB67296 Drosophil  |
| 52 | 7 | 3.9 | 186 | 4 | ABB66024 Drosophil  |
| 53 | 7 | 3.9 | 196 | 4 | ABG04750 Novel hum  |
| 54 | 7 | 3.9 | 212 | 3 | AAG55457 Arabidops  |
| 55 | 7 | 3.9 | 230 | 3 | AAG55456 Arabidops  |
| 56 | 7 | 3.9 | 233 | 6 | ABU23389 Protein e  |
| 57 | 7 | 3.9 | 238 | 4 | AAU43675 Propionib  |
| 58 | 7 | 3.9 | 238 | 6 | ABM40194 Propionib  |
| 59 | 7 | 3.9 | 248 | 3 | AAB17954 IL-1 anta  |
| 60 | 7 | 3.9 | 248 | 5 | ABR73422 Interleuk  |
| 61 | 7 | 3.9 | 258 | 4 | AAU58562 Propionib  |
| 62 | 7 | 3.9 | 258 | 6 | ABMS5081 Propionib  |
| 63 | 7 | 3.9 | 267 | 3 | AAG29599 Arabidops  |
| 64 | 7 | 3.9 | 271 | 3 | AAG29598 Arabidops  |
| 65 | 7 | 3.9 | 271 | 5 | ABBS0834 Herbicida  |
| 66 | 7 | 3.9 | 271 | 8 | ADN73459 Thale cre  |
| 67 | 7 | 3.9 | 285 | 3 | AAG29597 Arabidops  |
| 68 | 7 | 3.9 | 285 | 8 | ADO62704 Transcript |
| 69 | 7 | 3.9 | 303 | 8 | ADI43195 Plant tra  |
| 70 | 7 | 3.9 | 303 | 8 | ADO03266 Thalecres  |
| 71 | 7 | 3.9 | 328 | 4 | AAB60641 Moraxella  |
| 72 | 7 | 3.9 | 330 | 4 | ABG17524 Novel hum  |
| 73 | 7 | 3.9 | 335 | 2 | AAW92951 WO9905287  |
| 74 | 7 | 3.9 | 337 | 6 | ABM72575 Staphyloc  |
| 75 | 7 | 3.9 | 383 | 4 | AAU00428 Rat Gas1   |
| 76 | 7 | 3.9 | 405 | 6 | ABU18382 Protein e  |
| 77 | 7 | 3.9 | 409 | 6 | ABU18080 Protein e  |
| 78 | 7 | 3.9 | 418 | 7 | ABO76980 Pseudomon  |
| 79 | 7 | 3.9 | 419 | 6 | ABU00380 Human nov  |
| 80 | 7 | 3.9 | 456 | 7 | ADB70160 C. neofo   |
| 81 | 7 | 3.9 | 500 | 7 | ABO76964 Pseudomon  |
| 82 | 7 | 3.9 | 504 | 7 | ABO73545 Pseudomon  |
| 83 | 7 | 3.9 | 535 | 7 | ABO64944 Klebsiell  |
| 84 | 7 | 3.9 | 556 | 4 | ABBS62959 Drosophil |
| 85 | 7 | 3.9 | 558 | 5 | ABBS92528 Herbicida |
| 86 | 7 | 3.9 | 570 | 5 | ABBS1580 Herbicida  |
| 87 | 7 | 3.9 | 588 | 5 | AAW49551 Actinopla  |
| 88 | 7 | 3.9 | 601 | 7 | ADB70213 C. neofo   |
| 89 | 7 | 3.9 | 667 | 4 | ABBS66476 Drosophil |
| 90 | 7 | 3.9 | 678 | 4 | ABBS3847 Drosophil  |
| 91 | 7 | 3.9 | 688 | 2 | AAW43107 C. thermo  |
| 92 | 7 | 3.9 | 702 | 6 | ADA54694 Human pro  |
| 93 | 7 | 3.9 | 756 | 5 | ABBS49221 Listeria  |
| 94 | 7 | 3.9 | 814 | 2 | AAV11680 Sulfated   |
| 95 | 7 | 3.9 | 840 | 2 | AAW63112 Protein e  |
| 96 | 7 | 3.9 | 847 | 2 | AAW63108 Protein e  |
| 97 | 7 | 3.9 | 864 | 2 | AAU03636 Hypoxia-r  |
| 98 | 7 | 3.9 | 864 | 6 | ABU63755 Rat prote  |

|     |   |     |      |   |          |                    |     |   |     |    |   |          |                    |
|-----|---|-----|------|---|----------|--------------------|-----|---|-----|----|---|----------|--------------------|
| 99  | 7 | 3.9 | 864  | 7 | ADC69799 | Adc69799 Rat neuro | 172 | 6 | 3.3 | 32 | 2 | AAW82454 | Aaw82454 Onchocerc |
| 100 | 7 | 3.9 | 864  | 7 | ABW01151 | Abw01151 Hypoxia-r | 173 | 6 | 3.3 | 32 | 2 | AAW82452 | Aaw82452 X.gladius |
| 101 | 7 | 3.9 | 871  | 7 | ADB70269 | Abd70269 C. neofo  | 174 | 6 | 3.3 | 32 | 8 | ADO55707 | Ado55707 Swordfish |
| 102 | 7 | 3.9 | 887  | 4 | ABG03067 | Abg03067 Novel hum | 175 | 6 | 3.3 | 32 | 8 | ADO55709 | Ado55709 Onchocerc |
| 103 | 7 | 3.9 | 1015 | 7 | ADI39077 | Adi39077 Thermus t | 176 | 6 | 3.3 | 33 | 2 | AAW82457 | Aaw82457 S. cerevi |
| 104 | 7 | 3.9 | 1032 | 6 | AAg79679 | Ag79679 Human ENZ  | 177 | 6 | 3.3 | 33 | 8 | ADO55712 | Ado55712 Yeast cyt |
| 105 | 7 | 3.9 | 1048 | 6 | ABU36928 | Abu36928 Protein e | 178 | 6 | 3.3 | 34 | 3 | AAW56332 | Aaw56332 Human sec |
| 106 | 7 | 3.9 | 1056 | 4 | ABB70257 | Abb70257 Drosophil | 179 | 6 | 3.3 | 38 | 8 | ADJ72103 | Adj72103 T4 gene 3 |
| 107 | 7 | 3.9 | 1070 | 8 | ADO78099 | Ado78099 Human gua | 180 | 6 | 3.3 | 39 | 4 | AAW21845 | Aam21845 Peptide # |
| 108 | 7 | 3.9 | 1073 | 2 | AAW32063 | Aaw32063 Human ST  | 181 | 6 | 3.3 | 39 | 4 | AAW59878 | Aab59878 IGFBP-2 g |
| 109 | 7 | 3.9 | 1073 | 2 | AAW37371 | Aaw37371 Human ST  | 182 | 6 | 3.3 | 39 | 8 | ABO57730 | AbO57730 Human gen |
| 110 | 7 | 3.9 | 1073 | 5 | AAU08788 | Aau08788 Human gua | 183 | 6 | 3.3 | 39 | 8 | ABO57730 | AbO57730 Human gen |
| 111 | 7 | 3.9 | 1073 | 6 | ADA83736 | Ada83736 Human GUC | 184 | 6 | 3.3 | 40 | 2 | AAW11563 | Aaw11563 Human 5'  |
| 112 | 7 | 3.9 | 1073 | 8 | ADO78096 | Ado78096 Human gua | 185 | 6 | 3.3 | 43 | 3 | AAW76369 | Aay76369 Fragment  |
| 113 | 7 | 3.9 | 1075 | 2 | AAW38861 | Aar38861 GC-C, 2/1 | 186 | 6 | 3.3 | 43 | 3 | AAW87166 | Aam87166 Human imm |
| 114 | 7 | 3.9 | 1232 | 6 | ABU24270 | Abu24270 Protein e | 187 | 6 | 3.3 | 43 | 7 | ADE12017 | Ade12017 Human sec |
| 115 | 7 | 3.9 | 1240 | 5 | ABP29526 | Abp29526 Streptoco | 188 | 6 | 3.3 | 46 | 4 | AAW63399 | Aab63399 Human bre |
| 116 | 7 | 3.9 | 1323 | 8 | ADO55192 | Ado55192 Protein # | 189 | 6 | 3.3 | 47 | 4 | ADG27718 | Adg27718 Human nov |
| 117 | 7 | 3.9 | 1325 | 8 | ADO55191 | Ado55191 Protein # | 190 | 6 | 3.3 | 48 | 4 | ABB23040 | Abb23040 Protein # |
| 118 | 7 | 3.9 | 1327 | 3 | AAW70474 | Aay70474 Human cyc | 191 | 6 | 3.3 | 51 | 3 | AAW38570 | Aab38570 Gene 2 hu |
| 119 | 7 | 3.9 | 1389 | 4 | ABB58963 | Abb58963 Drosophil | 192 | 6 | 3.3 | 52 | 3 | AAW53979 | Aab53979 Human col |
| 120 | 7 | 3.9 | 1487 | 6 | ABU49518 | Abu49518 Protein e | 193 | 6 | 3.3 | 52 | 5 | AAE17854 | Aae17854 Sequence3 |
| 121 | 7 | 3.9 | 1679 | 6 | ABR52863 | AbR52863 Protein s | 194 | 6 | 3.3 | 53 | 5 | AAE17853 | Aae17853 Sequence2 |
| 122 | 7 | 3.9 | 1679 | 7 | ADK62332 | Adk62332 Disease t | 195 | 6 | 3.3 | 54 | 4 | AAU45079 | Aau45079 Propionib |
| 123 | 6 | 3.3 | 9    | 2 | AAW40616 | Aay40616 Al deriva | 196 | 6 | 3.3 | 54 | 6 | AAW41598 | Aam41598 Propionib |
| 124 | 6 | 3.3 | 9    | 3 | AAW29955 | Aab29955 Scaffold  | 197 | 6 | 3.3 | 56 | 4 | AAW87923 | Aam87923 Human imm |
| 125 | 6 | 3.3 | 9    | 6 | AAW35060 | Aae35060 Immunoglo | 198 | 6 | 3.3 | 56 | 4 | ABB17912 | Abb17912 Human ner |
| 126 | 6 | 3.3 | 9    | 6 | AAW35075 | Aae35075 Human imm | 199 | 6 | 3.3 | 57 | 4 | ABB17544 | Abb17544 Human ner |
| 127 | 6 | 3.3 | 13   | 5 | ABG92792 | Abg92792 C. carbon | 200 | 6 | 3.3 | 57 | 5 | ABP29616 | Abp29616 Streptoco |
| 128 | 6 | 3.3 | 13   | 6 | ABR75740 | AbR75740 Liver res | 201 | 6 | 3.3 | 57 | 8 | ABO54552 | AbO54552 Human gen |
| 129 | 6 | 3.3 | 13   | 7 | ADN07530 | Adn07530 Liver res | 202 | 6 | 3.3 | 58 | 4 | AAW89124 | Aaw89124 Human imm |
| 130 | 6 | 3.3 | 18   | 2 | AAW18019 | Aaw18019 Human 77  | 203 | 6 | 3.3 | 58 | 4 | AAU64960 | Aau64960 Propionib |
| 131 | 6 | 3.3 | 20   | 4 | AAE12293 | Aae12293 Mycobacte | 204 | 6 | 3.3 | 58 | 6 | ABM61479 | Abm61479 Propionib |
| 132 | 6 | 3.3 | 20   | 7 | ADL34420 | Adl34420 MHC/HLA p | 205 | 6 | 3.3 | 59 | 4 | AAU64696 | Aau64696 Propionib |
| 133 | 6 | 3.3 | 21   | 2 | AAW99879 | Aar99879 Thrombin  | 206 | 6 | 3.3 | 59 | 6 | ABM43015 | Abm43015 Propionib |
| 134 | 6 | 3.3 | 21   | 4 | AAW20462 | Aam20462 Peptide # | 207 | 6 | 3.3 | 60 | 4 | AAU41199 | Aau41199 Propionib |
| 135 | 6 | 3.3 | 21   | 4 | ABW41290 | Abw41290 Peptide # | 208 | 6 | 3.3 | 60 | 6 | ABM37718 | Abm37718 Propionib |
| 136 | 6 | 3.3 | 21   | 4 | AAW35075 | Aam35075 Peptide # | 209 | 6 | 3.3 | 61 | 5 | ADP35550 | Adp35550 Human ORF |
| 137 | 6 | 3.3 | 21   | 4 | ABB25262 | Abb25262 Peptide # | 210 | 6 | 3.3 | 63 | 6 | ADA34693 | Ada34693 Acinetoba |
| 138 | 6 | 3.3 | 21   | 4 | AAW74959 | Aam74959 Human bon | 211 | 6 | 3.3 | 63 | 4 | AAU58837 | Aau58837 Propionib |
| 139 | 6 | 3.3 | 21   | 4 | AAW62155 | Aam62155 Human bra | 212 | 6 | 3.3 | 63 | 4 | AAU65853 | Aau65853 Propionib |
| 140 | 6 | 3.3 | 21   | 4 | ABG56732 | Abg56732 Human liv | 213 | 6 | 3.3 | 63 | 4 | AAU44236 | Aau44236 Propionib |
| 141 | 6 | 3.3 | 21   | 5 | ABG44698 | Abg44698 Human pep | 214 | 6 | 3.3 | 63 | 6 | ABM55356 | Abm55356 Propionib |
| 142 | 6 | 3.3 | 22   | 2 | AAW99884 | Aar99884 Thrombin  | 215 | 6 | 3.3 | 63 | 6 | ABM62372 | Abm62372 Propionib |
| 143 | 6 | 3.3 | 24   | 6 | ABJ38272 | Abj38272 TALI-1 re | 216 | 6 | 3.3 | 63 | 6 | ABM40755 | Abm40755 Propionib |
| 144 | 6 | 3.3 | 24   | 6 | AAO27002 | Aao27002 Human Ngr | 217 | 6 | 3.3 | 65 | 3 | AAW03505 | Aaw03505 Human sec |
| 145 | 6 | 3.3 | 25   | 3 | AAW81885 | Aay81885 Yeast SOD | 218 | 6 | 3.3 | 65 | 4 | AAO07090 | Aao07090 Human pol |
| 146 | 6 | 3.3 | 26   | 4 | AAW15756 | Aam15756 Peptide # | 219 | 6 | 3.3 | 66 | 3 | AAW43817 | Aag43817 Arabidops |
| 147 | 6 | 3.3 | 26   | 4 | AAW16276 | Aam16276 Peptide # | 220 | 6 | 3.3 | 67 | 3 | AAW56331 | Aab56331 Human sec |
| 148 | 6 | 3.3 | 26   | 4 | ABB34753 | Abb34753 Peptide # | 221 | 6 | 3.3 | 67 | 4 | ABG00995 | Abg00995 Novel hum |
| 149 | 6 | 3.3 | 26   | 4 | ABB35264 | Abb35264 Peptide # | 222 | 6 | 3.3 | 68 | 3 | AAW33178 | Aag33178 Zea may   |
| 150 | 6 | 3.3 | 26   | 4 | AAW28762 | Aam28762 Peptide # | 223 | 6 | 3.3 | 68 | 5 | ABB49671 | Abb49671 Listeria  |
| 151 | 6 | 3.3 | 26   | 4 | AAW28265 | Aam28265 Peptide # | 224 | 6 | 3.3 | 69 | 4 | AAO08961 | Aao08961 Human pol |
| 152 | 6 | 3.3 | 26   | 4 | ABB30093 | Abb30093 Peptide # | 225 | 6 | 3.3 | 70 | 4 | AAU53679 | Aau53679 Propionib |
| 153 | 6 | 3.3 | 26   | 4 | ABB29576 | Abb29576 Peptide # | 226 | 6 | 3.3 | 70 | 5 | ABJ05443 | Abj05443 Human bre |
| 154 | 6 | 3.3 | 26   | 4 | ABB20168 | Abb20168 Protein # | 227 | 6 | 3.3 | 70 | 6 | ABM50198 | Abm50198 Propionib |
| 155 | 6 | 3.3 | 26   | 4 | ABB20706 | Abb20706 Protein # | 228 | 6 | 3.3 | 71 | 4 | AAO09580 | Aao09580 Human pol |
| 156 | 6 | 3.3 | 26   | 4 | AAW68465 | Aam68465 Human bon | 229 | 6 | 3.3 | 72 | 7 | ADC97580 | Adc97580 E. faeciu |
| 157 | 6 | 3.3 | 26   | 4 | AAW67939 | Aam67939 Human bon | 230 | 6 | 3.3 | 73 | 3 | AAW52931 | Aag52931 Arabidops |
| 158 | 6 | 3.3 | 26   | 4 | AAW56095 | Aam56095 Human bra | 231 | 6 | 3.3 | 73 | 3 | AAW07581 | Aag07581 Arabidops |
| 159 | 6 | 3.3 | 26   | 4 | AAW55555 | Aam55555 Human bra | 232 | 6 | 3.3 | 73 | 3 | AAW36961 | Aag36961 Arabidops |
| 160 | 6 | 3.3 | 26   | 4 | ABG49581 | Abg49581 Human liv | 233 | 6 | 3.3 | 73 | 3 | AAW05446 | Aag05446 Arabidops |
| 161 | 6 | 3.3 | 26   | 4 | ABG50138 | Abg50138 Human liv | 234 | 6 | 3.3 | 73 | 3 | AAW08766 | Aag08766 Arabidops |
| 162 | 6 | 3.3 | 26   | 4 | AAW03490 | Aam03490 Peptide # | 235 | 6 | 3.3 | 73 | 3 | AAW08148 | Aag08148 Arabidops |
| 163 | 6 | 3.3 | 26   | 4 | AAW04009 | Aam04009 Peptide # | 236 | 6 | 3.3 | 73 | 3 | AAW36038 | Aag36038 Zea may   |
| 164 | 6 | 3.3 | 26   | 5 | ABG38046 | Abg38046 Human pep | 237 | 6 | 3.3 | 73 | 4 | AAW19295 | Aam19295 Peptide # |
| 165 | 6 | 3.3 | 26   | 5 | ABG37473 | Abg37473 Human pep | 238 | 6 | 3.3 | 73 | 4 | ABW38577 | Abw38577 Peptide # |
| 166 | 6 | 3.3 | 27   | 7 | ADJ81155 | Adj81155 Self-asse | 239 | 6 | 3.3 | 73 | 4 | AAW32028 | Aam32028 Peptide # |
| 167 | 6 | 3.3 | 27   | 7 | ADJ81362 | Adj81362 Self-asse | 240 | 6 | 3.3 | 73 | 4 | ABB23701 | Abb23701 Human bon |
| 168 | 6 | 3.3 | 27   | 7 | ADJ81232 | Adj81232 Self-asse | 241 | 6 | 3.3 | 73 | 4 | AAW71736 | Aam71736 Human bra |
| 169 | 6 | 3.3 | 27   | 7 | ADJ81293 | Adj81293 Self-asse | 242 | 6 | 3.3 | 73 | 4 | AAW59201 | Aam59201 Human liv |
| 170 | 6 | 3.3 | 29   | 7 | ADK40717 | Adk40717 kDR & VEG | 243 | 6 | 3.3 | 73 | 4 | ABG53421 | Abg53421 Human pep |
| 171 | 6 | 3.3 | 29   | 7 | ADK40695 | Adk40695 KDR & VEG | 244 | 6 | 3.3 | 73 | 5 | ABG41550 | Abg41550 Human pep |

|     |   |      |     |   |          |           |     |   |      |     |   |          |           |
|-----|---|------|-----|---|----------|-----------|-----|---|------|-----|---|----------|-----------|
| 245 | 6 | 3..3 | 76  | 2 | AA43264  | Human pat | 318 | 6 | 3..3 | 112 | 7 | ADE46275 | Human car |
| 246 | 6 | 3..3 | 76  | 4 | ABM84006 | Human imm | 319 | 6 | 3..3 | 113 | 5 | ABG14880 | Novel hum |
| 247 | 6 | 3..3 | 76  | 6 | ABG74105 | Human pat | 320 | 6 | 3..3 | 113 | 5 | ABG99297 | Programme |
| 248 | 6 | 3..3 | 76  | 8 | ABO60237 | Human gen | 321 | 6 | 3..3 | 113 | 6 | ADA54961 | Human pro |
| 249 | 6 | 3..3 | 78  | 4 | AAU49415 | Propionib | 322 | 6 | 3..3 | 113 | 6 | ABJ26492 | Aspergill |
| 250 | 6 | 3..3 | 78  | 4 | AAU20719 | Human nov | 323 | 6 | 3..3 | 113 | 7 | ADM04171 | Human pro |
| 251 | 6 | 3..3 | 78  | 6 | ABM45934 | Propionib | 324 | 6 | 3..3 | 114 | 4 | AAO03076 | Human pol |
| 252 | 6 | 3..3 | 79  | 3 | ABP40984 | Human ret | 325 | 6 | 3..3 | 114 | 4 | AAU53097 | Propionib |
| 253 | 6 | 3..3 | 79  | 4 | AAU14499 | Human nov | 326 | 6 | 3..3 | 114 | 6 | ABM49616 | Propionib |
| 254 | 6 | 3..3 | 79  | 8 | ADH80817 | Human pol | 327 | 6 | 3..3 | 115 | 5 | ABU51777 | Helicobac |
| 255 | 6 | 3..3 | 80  | 5 | ABP66679 | Human bre | 328 | 6 | 3..3 | 116 | 4 | AAO03015 | Human pol |
| 256 | 6 | 3..3 | 81  | 4 | ABU52845 | Propionib | 329 | 6 | 3..3 | 116 | 4 | ABG21772 | Novel hum |
| 257 | 6 | 3..3 | 81  | 6 | ABM49364 | Propionib | 330 | 6 | 3..3 | 116 | 6 | ABP79492 | N. gonorr |
| 258 | 6 | 3..3 | 82  | 4 | AAU14263 | Human nov | 331 | 6 | 3..3 | 117 | 4 | AAO00804 | Human pol |
| 259 | 6 | 3..3 | 84  | 4 | ABG12684 | Human nov | 332 | 6 | 3..3 | 117 | 4 | AAO03690 | Human pol |
| 260 | 6 | 3..3 | 85  | 3 | AAU33390 | Novel hum | 333 | 6 | 3..3 | 118 | 3 | ABM42475 | Human ORF |
| 261 | 6 | 3..3 | 86  | 4 | AAU18310 | Peptide # | 334 | 6 | 3..3 | 119 | 4 | ABG15528 | Novel hum |
| 262 | 6 | 3..3 | 86  | 4 | AAU70473 | Human bon | 335 | 6 | 3..3 | 120 | 2 | AAW52824 | Human TFE |
| 263 | 6 | 3..3 | 86  | 4 | AAU05916 | Peptide # | 336 | 6 | 3..3 | 120 | 5 | ABG77273 | Selected  |
| 264 | 6 | 3..3 | 86  | 5 | ABG40112 | Human pep | 337 | 6 | 3..3 | 120 | 5 | ABJ11144 | Yeast sel |
| 265 | 6 | 3..3 | 86  | 7 | ADD27592 | Human adi | 338 | 6 | 3..3 | 121 | 2 | AAW95647 | Mus muscu |
| 266 | 6 | 3..3 | 87  | 5 | ABR01723 | Human bre | 339 | 6 | 3..3 | 121 | 2 | AAW14872 | Peptide # |
| 267 | 6 | 3..3 | 88  | 4 | AAO03796 | Human pol | 340 | 6 | 3..3 | 121 | 4 | ABM33840 | Peptide # |
| 268 | 6 | 3..3 | 88  | 4 | AAU51153 | Propionib | 341 | 6 | 3..3 | 121 | 4 | ABM27299 | Peptide # |
| 269 | 6 | 3..3 | 88  | 6 | ABM47672 | Propionib | 342 | 6 | 3..3 | 121 | 4 | ABM28656 | Peptide # |
| 270 | 6 | 3..3 | 89  | 4 | AAO07402 | Human pol | 343 | 6 | 3..3 | 121 | 4 | ABM19282 | Protein # |
| 271 | 6 | 3..3 | 92  | 3 | AAU26179 | Zea mays  | 344 | 6 | 3..3 | 121 | 4 | AAW67012 | Human bon |
| 272 | 6 | 3..3 | 93  | 3 | AAU48666 | Arabidops | 345 | 6 | 3..3 | 121 | 4 | AAW54606 | Human bra |
| 273 | 6 | 3..3 | 93  | 5 | ABJ10323 | Human hum | 346 | 6 | 3..3 | 121 | 4 | AAW76936 | Variable  |
| 274 | 6 | 3..3 | 93  | 5 | ADK35105 | Novel hum | 347 | 6 | 3..3 | 121 | 4 | ABG48675 | Human liv |
| 275 | 6 | 3..3 | 93  | 7 | ADF18288 | Insulin-1 | 348 | 6 | 3..3 | 121 | 4 | AAW02597 | Peptide # |
| 276 | 6 | 3..3 | 94  | 3 | AAU44476 | Arabidops | 349 | 6 | 3..3 | 121 | 5 | ABG36669 | Human pep |
| 277 | 6 | 3..3 | 95  | 4 | AAO00887 | Human bon | 350 | 6 | 3..3 | 121 | 8 | ADN07023 | Murine an |
| 278 | 6 | 3..3 | 95  | 4 | ABG27776 | Novel hum | 351 | 6 | 3..3 | 122 | 4 | AAO06277 | Human pol |
| 279 | 6 | 3..3 | 95  | 6 | ABM65682 | Propionib | 352 | 6 | 3..3 | 123 | 4 | AAW23580 | Murine ES |
| 280 | 6 | 3..3 | 95  | 6 | ABJ25892 | Aspergill | 353 | 6 | 3..3 | 123 | 4 | AAO06597 | Human pol |
| 281 | 6 | 3..3 | 96  | 6 | ABM65153 | Propionib | 354 | 6 | 3..3 | 123 | 4 | ABM17908 | Human ner |
| 282 | 6 | 3..3 | 98  | 5 | ADK34204 | Novel hum | 355 | 6 | 3..3 | 124 | 3 | AAU26740 | Zea mays  |
| 283 | 6 | 3..3 | 99  | 4 | AAU53964 | Propionib | 356 | 6 | 3..3 | 125 | 3 | AAU36408 | Arabidops |
| 284 | 6 | 3..3 | 99  | 6 | ABM50483 | Propionib | 357 | 6 | 3..3 | 125 | 4 | AAU67165 | Propionib |
| 285 | 6 | 3..3 | 101 | 4 | AAO07556 | Human pol | 358 | 6 | 3..3 | 125 | 5 | ABP09553 | Human ORF |
| 286 | 6 | 3..3 | 102 | 4 | AAU40929 | Propionib | 359 | 6 | 3..3 | 125 | 7 | ADM04184 | Human pro |
| 287 | 6 | 3..3 | 102 | 6 | ABM37448 | Propionib | 360 | 6 | 3..3 | 126 | 3 | AAU19630 | Arabidops |
| 288 | 6 | 3..3 | 103 | 3 | AAU65095 | Human 5'  | 361 | 6 | 3..3 | 127 | 5 | ABM49494 | Listeria  |
| 289 | 6 | 3..3 | 103 | 4 | AAU46895 | Propionib | 362 | 6 | 3..3 | 127 | 5 | ABU70811 | Human adi |
| 290 | 6 | 3..3 | 103 | 6 | ABM43414 | Propionib | 363 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 291 | 6 | 3..3 | 105 | 3 | AAU47232 | Arabidops | 364 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 292 | 6 | 3..3 | 105 | 3 | AAU20186 | Arabidops | 365 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 293 | 6 | 3..3 | 105 | 5 | AAO22559 | Wooden le | 366 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 294 | 6 | 3..3 | 105 | 7 | ADC96946 | E. faeciu | 367 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 295 | 6 | 3..3 | 105 | 7 | ADF05795 | Bacterial | 368 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 296 | 6 | 3..3 | 106 | 3 | AAU05445 | Arabidops | 369 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 297 | 6 | 3..3 | 106 | 3 | AAU52930 | Arabidops | 370 | 6 | 3..3 | 128 | 3 | AAU18964 | Zea mays  |
| 298 | 6 | 3..3 | 106 | 3 | AAU08765 | Arabidops | 371 | 6 | 3..3 | 128 | 5 | AAU45066 | Propionib |
| 299 | 6 | 3..3 | 107 | 3 | ABM40288 | Human ORF | 372 | 6 | 3..3 | 128 | 6 | ABM41585 | Propionib |
| 300 | 6 | 3..3 | 107 | 5 | ABP00707 | Human ORF | 373 | 6 | 3..3 | 129 | 4 | AAU1585  | Propionib |
| 301 | 6 | 3..3 | 108 | 5 | ABP38241 | Staphyloc | 374 | 6 | 3..3 | 130 | 4 | AAU17522 | Peptide # |
| 302 | 6 | 3..3 | 108 | 6 | ABP56102 | Human IGP | 375 | 6 | 3..3 | 130 | 4 | ABM36543 | Peptide # |
| 303 | 6 | 3..3 | 109 | 3 | AAU28188 | Human car | 376 | 6 | 3..3 | 130 | 4 | AAU30043 | Peptide # |
| 304 | 6 | 3..3 | 109 | 3 | AAU36036 | Zea mays  | 377 | 6 | 3..3 | 130 | 4 | ABM31337 | Peptide # |
| 305 | 6 | 3..3 | 109 | 7 | ADC95372 | E. faeciu | 378 | 6 | 3..3 | 130 | 4 | AAU69709 | Human bon |
| 306 | 6 | 3..3 | 109 | 7 | ADC96779 | E. faeciu | 379 | 6 | 3..3 | 130 | 4 | AAU69709 | Human bon |
| 307 | 6 | 3..3 | 110 | 3 | AAU26178 | Zea mays  | 380 | 6 | 3..3 | 130 | 5 | AAU69709 | Human bon |
| 308 | 6 | 3..3 | 110 | 6 | ABR54947 | IGG light | 381 | 6 | 3..3 | 131 | 4 | AAU69709 | Human bon |
| 309 | 6 | 3..3 | 110 | 6 | ADA36586 | Acinetoba | 382 | 6 | 3..3 | 131 | 4 | AAU69709 | Human bon |
| 310 | 6 | 3..3 | 111 | 4 | AAO06659 | Human pol | 383 | 6 | 3..3 | 131 | 6 | ABM65276 | Propionib |
| 311 | 6 | 3..3 | 111 | 4 | AAO03548 | Human pol | 384 | 6 | 3..3 | 132 | 3 | ABM65276 | Propionib |
| 312 | 6 | 3..3 | 111 | 4 | ABM14932 | Human ner | 385 | 6 | 3..3 | 132 | 4 | AAO05115 | Human pol |
| 313 | 6 | 3..3 | 111 | 7 | ADH86899 | Enterococ | 386 | 6 | 3..3 | 132 | 4 | ABG00130 | Novel hum |
| 314 | 6 | 3..3 | 112 | 3 | AAU51825 | Human sec | 387 | 6 | 3..3 | 133 | 2 | ABP06906 | Human ORF |
| 315 | 6 | 3..3 | 112 | 4 | AAU22307 | Human car | 388 | 6 | 3..3 | 133 | 2 | ABM38183 | Arabidops |
| 316 | 6 | 3..3 | 112 | 6 | ABR54943 | IGG light | 389 | 6 | 3..3 | 134 | 2 | ABM65250 | Propionib |
| 317 | 6 | 3..3 | 112 | 6 | ABR54938 | IGG light | 390 | 6 | 3..3 | 134 | 3 | AAU85194 | Heavy cha |

|     |   |     |     |   |           |                     |     |   |          |                    |
|-----|---|-----|-----|---|-----------|---------------------|-----|---|----------|--------------------|
| 391 | 6 | 3.3 | 134 | 3 | AAG03782  | Aag03782 Human sec  | 464 | 1 | AAP91963 | Aap91963 MPB-70 pr |
| 392 | 6 | 3.3 | 134 | 3 | AAG20185  | Aag20185 Arabidops  | 465 | 2 | AAR07053 | Aar07053 Immunopro |
| 393 | 6 | 3.3 | 134 | 3 | AAQ47231  | Aag47231 Arabidops  | 466 | 3 | AAW12045 | Aaw12045 MPB 70. 4 |
| 394 | 6 | 3.3 | 135 | 4 | AAO13042  | Aao13042 Human pol  | 467 | 6 | ABM72205 | Abm72205 Staphyloc |
| 395 | 6 | 3.3 | 136 | 4 | AAE00629  | Aae00629 Rice NADH  | 468 | 6 | ABb49803 | Abb49803 Listeria  |
| 396 | 6 | 3.3 | 136 | 4 | AAU29701  | Aau29701 Novel hum  | 469 | 6 | ABu05627 | Abu05627 M. tuberc |
| 397 | 6 | 3.3 | 137 | 4 | ABG10351  | Abg10351 Novel hum  | 470 | 6 | ABP32114 | Abp32114 Human ORF |
| 398 | 6 | 3.3 | 137 | 4 | AAQ19629  | Aag19629 Arabidops  | 471 | 6 | ABU26428 | Abu26428 Protein e |
| 399 | 6 | 3.3 | 138 | 3 | AAQ32263  | Aag32263 Arabidops  | 472 | 6 | ABU51918 | Abu51918 Helicobac |
| 400 | 6 | 3.3 | 138 | 4 | AAQ20095  | Aag20095 Novel hum  | 473 | 6 | ABP62924 | Abp62924 Human pol |
| 401 | 6 | 3.3 | 139 | 3 | AAQ33389  | Aag33389 Zea mays   | 474 | 6 | ABP65755 | Abp65755 Human pro |
| 402 | 6 | 3.3 | 139 | 4 | AAQ86378  | Aam86378 Human imm  | 475 | 6 | AAy96467 | Aay96467 Partial V |
| 403 | 6 | 3.3 | 139 | 5 | ABP40071  | Abp40071 Staphyloc  | 476 | 6 | AAy96467 | Aay96467 Helicobac |
| 404 | 6 | 3.3 | 141 | 3 | AAQ15009  | Aag15009 Arabidops  | 477 | 6 | AAy96467 | Aay96467 Helicobac |
| 405 | 6 | 3.3 | 141 | 4 | AAO13630  | Aao13630 Human pol  | 478 | 6 | AAy96467 | Aay96467 Helicobac |
| 406 | 6 | 3.3 | 141 | 4 | AAQ05534  | Aag05534 Novel hum  | 479 | 6 | AAy96467 | Aay96467 Helicobac |
| 407 | 6 | 3.3 | 141 | 6 | ABP58742  | Abp58742 Human reg  | 480 | 6 | AAy96467 | Aay96467 Helicobac |
| 408 | 6 | 3.3 | 141 | 6 | ABU00666  | Abu00666 Human nov  | 481 | 6 | AAy96467 | Aay96467 Helicobac |
| 409 | 6 | 3.3 | 141 | 7 | ADC33248  | Adc33248 Human nov  | 482 | 6 | AAy96467 | Aay96467 Helicobac |
| 410 | 6 | 3.3 | 142 | 7 | ABO61865  | Abo61865 Klebsiell  | 483 | 6 | AAy96467 | Aay96467 Helicobac |
| 411 | 6 | 3.3 | 143 | 4 | AAU56751  | Aau56751 Propionib  | 484 | 6 | AAy96467 | Aay96467 Helicobac |
| 412 | 6 | 3.3 | 143 | 4 | ABR96190  | AbR96190 Human NOV  | 485 | 6 | AAy96467 | Aay96467 Helicobac |
| 413 | 6 | 3.3 | 143 | 6 | ABR96190  | AbR96190 Human NOV  | 486 | 6 | AAy96467 | Aay96467 Helicobac |
| 414 | 6 | 3.3 | 143 | 6 | ABR96190  | AbR96190 Human NOV  | 487 | 6 | AAy96467 | Aay96467 Helicobac |
| 415 | 6 | 3.3 | 143 | 7 | ADM26138  | Adm26138 Hyperther  | 488 | 6 | AAy96467 | Aay96467 Helicobac |
| 416 | 6 | 3.3 | 145 | 4 | AAQ86678  | Aam86678 Human imm  | 489 | 6 | AAy96467 | Aay96467 Helicobac |
| 417 | 6 | 3.3 | 145 | 4 | ABR11556  | Abb11556 Human bre  | 490 | 6 | AAy96467 | Aay96467 Helicobac |
| 418 | 6 | 3.3 | 146 | 4 | AAQ43505  | Aam43505 Human pol  | 491 | 6 | AAy96467 | Aay96467 Helicobac |
| 419 | 6 | 3.3 | 146 | 4 | AAQ229314 | Abg229314 Novel hum | 492 | 6 | AAy96467 | Aay96467 Helicobac |
| 420 | 6 | 3.3 | 146 | 6 | ABU36050  | Abu36050 Protein e  | 493 | 6 | AAy96467 | Aay96467 Helicobac |
| 421 | 6 | 3.3 | 146 | 8 | ADM24526  | Adm24526 Human PRO  | 494 | 6 | AAy96467 | Aay96467 Helicobac |
| 422 | 6 | 3.3 | 147 | 4 | AAQ93730  | Aau93730 Human pol  | 495 | 6 | AAy96467 | Aay96467 Helicobac |
| 423 | 6 | 3.3 | 147 | 4 | AAU42003  | Aau42003 Propionib  | 496 | 6 | AAy96467 | Aay96467 Helicobac |
| 424 | 6 | 3.3 | 147 | 6 | ABR38522  | Abu38522 Propionib  | 497 | 6 | AAy96467 | Aay96467 Helicobac |
| 425 | 6 | 3.3 | 147 | 6 | ABU33940  | Abu33940 Protein e  | 498 | 6 | AAy96467 | Aay96467 Helicobac |
| 426 | 6 | 3.3 | 147 | 8 | ADL31658  | Adl31658 Human pro  | 499 | 6 | AAy96467 | Aay96467 Helicobac |
| 427 | 6 | 3.3 | 148 | 4 | AAO05143  | Aao05143 Human pol  | 500 | 6 | AAy96467 | Aay96467 Helicobac |
| 428 | 6 | 3.3 | 148 | 4 | AAQ95353  | Aab95353 Human pro  | 501 | 6 | AAy96467 | Aay96467 Helicobac |
| 429 | 6 | 3.3 | 149 | 5 | AAQ52488  | Aam52488 Superoxid  | 502 | 6 | AAy96467 | Aay96467 Helicobac |
| 430 | 6 | 3.3 | 149 | 6 | ADA06217  | Ada06217 Rice chol  | 503 | 6 | AAy96467 | Aay96467 Helicobac |
| 431 | 6 | 3.3 | 149 | 7 | ADC33173  | Adc33173 Human nov  | 504 | 6 | AAy96467 | Aay96467 Helicobac |
| 432 | 6 | 3.3 | 149 | 8 | ADO16988  | Ado16988 Rice chol  | 505 | 6 | AAy96467 | Aay96467 Helicobac |
| 433 | 6 | 3.3 | 150 | 5 | AAQ52487  | Aam52487 Superoxid  | 506 | 6 | AAy96467 | Aay96467 Helicobac |
| 434 | 6 | 3.3 | 151 | 3 | AAQ36407  | Aag36407 Arabidops  | 507 | 6 | AAy96467 | Aay96467 Helicobac |
| 435 | 6 | 3.3 | 152 | 2 | AAQ56672  | Aaw56672 Paddy cop  | 508 | 6 | AAy96467 | Aay96467 Helicobac |
| 436 | 6 | 3.3 | 152 | 3 | ABR12538  | Abi12538 Rice supe  | 509 | 6 | AAy96467 | Aay96467 Helicobac |
| 437 | 6 | 3.3 | 152 | 7 | ADB85176  | Adb85176 Rat supe   | 510 | 6 | AAy96467 | Aay96467 Helicobac |
| 438 | 6 | 3.3 | 153 | 1 | AAQ70979  | Aap70979 Metalfree  | 511 | 6 | AAy96467 | Aay96467 Helicobac |
| 439 | 6 | 3.3 | 153 | 3 | AAQ15008  | Aag15008 Arabidops  | 512 | 6 | AAy96467 | Aay96467 Helicobac |
| 440 | 6 | 3.3 | 153 | 5 | AAQ52490  | Aam52490 Superoxid  | 513 | 6 | AAy96467 | Aay96467 Helicobac |
| 441 | 6 | 3.3 | 153 | 5 | AAQ52490  | Aam52490 Superoxid  | 514 | 6 | AAy96467 | Aay96467 Helicobac |
| 442 | 6 | 3.3 | 153 | 7 | ADD48315  | Ad48315 Rat prote   | 515 | 6 | AAy96467 | Aay96467 Helicobac |
| 443 | 6 | 3.3 | 154 | 5 | AAQ52499  | Ad48679 Rat prote   | 516 | 6 | AAy96467 | Aay96467 Helicobac |
| 444 | 6 | 3.3 | 154 | 5 | AAQ52499  | Ad48679 Rat prote   | 517 | 6 | AAy96467 | Aay96467 Helicobac |
| 445 | 6 | 3.3 | 154 | 5 | AAQ52499  | Ad48679 Rat prote   | 518 | 6 | AAy96467 | Aay96467 Helicobac |
| 446 | 6 | 3.3 | 155 | 3 | AAQ17607  | Aag17607 Arabidops  | 519 | 6 | AAy96467 | Aay96467 Helicobac |
| 447 | 6 | 3.3 | 155 | 5 | ADK36829  | Adk36829 Novel hum  | 520 | 6 | AAy96467 | Aay96467 Helicobac |
| 448 | 6 | 3.3 | 156 | 7 | ABO64991  | Abo64991 Klebsiell  | 521 | 6 | AAy96467 | Aay96467 Helicobac |
| 449 | 6 | 3.3 | 156 | 8 | ADO05624  | Ado05624 Human EXM  | 522 | 6 | AAy96467 | Aay96467 Helicobac |
| 450 | 6 | 3.3 | 157 | 3 | AAQ17606  | Aag17606 Arabidops  | 523 | 6 | AAy96467 | Aay96467 Helicobac |
| 451 | 6 | 3.3 | 157 | 3 | AAQ17606  | Aag17606 Arabidops  | 524 | 6 | AAy96467 | Aay96467 Helicobac |
| 452 | 6 | 3.3 | 157 | 6 | ABM44272  | Abm44272 Propionib  | 525 | 6 | AAy96467 | Aay96467 Helicobac |
| 453 | 6 | 3.3 | 157 | 7 | ABO75105  | Abo75105 Pseudomon  | 526 | 6 | AAy96467 | Aay96467 Helicobac |
| 454 | 6 | 3.3 | 159 | 3 | AAQ54341  | Aay54341 Amino aci  | 527 | 6 | AAy96467 | Aay96467 Helicobac |
| 455 | 6 | 3.3 | 159 | 6 | AAU43632  | Aau43632 Propionib  | 528 | 6 | AAy96467 | Aay96467 Helicobac |
| 456 | 6 | 3.3 | 159 | 6 | ABM40151  | Abm40151 Propionib  | 529 | 6 | AAy96467 | Aay96467 Helicobac |
| 457 | 6 | 3.3 | 161 | 3 | AAQ26177  | Aag26177 Zea mays   | 530 | 6 | AAy96467 | Aay96467 Helicobac |
| 458 | 6 | 3.3 | 161 | 5 | ABU48463  | Abu48463 Listeria   | 531 | 6 | AAy96467 | Aay96467 Helicobac |
| 459 | 6 | 3.3 | 161 | 6 | ABU34684  | Abu34684 Protein e  | 532 | 6 | AAy96467 | Aay96467 Helicobac |
| 460 | 6 | 3.3 | 161 | 6 | ABU34734  | Abu34734 Protein e  | 533 | 6 | AAy96467 | Aay96467 Helicobac |
| 461 | 6 | 3.3 | 162 | 3 | AAQ32262  | Aag32262 Arabidops  | 534 | 6 | AAy96467 | Aay96467 Helicobac |
| 462 | 6 | 3.3 | 162 | 4 | AAU59894  | Aau59894 Propionib  | 535 | 6 | AAy96467 | Aay96467 Helicobac |
| 463 | 6 | 3.3 | 162 | 6 | ABM56413  | Abm56413 Propionib  | 536 | 6 | AAy96467 | Aay96467 Helicobac |
|     |   |     | 162 | 7 | ADFI13844 | Adfi13844 Human end |     |   |          |                    |

|     |   |     |     |   |          |                     |     |   |     |     |   |          |                    |
|-----|---|-----|-----|---|----------|---------------------|-----|---|-----|-----|---|----------|--------------------|
| 537 | 6 | 3.3 | 195 | 4 | ABB21123 | Abb21123 Protein #  | 610 | 6 | 3.3 | 208 | 6 | ABU20824 | Abu20824 Protein e |
| 538 | 6 | 3.3 | 195 | 4 | ABG50547 | Abg50547 Human liv  | 611 | 6 | 3.3 | 208 | 6 | AAO24019 | Aao24019 Human hep |
| 539 | 6 | 3.3 | 195 | 5 | AAM52477 | Aam52477 Superoxid  | 612 | 6 | 3.3 | 208 | 7 | ADC34571 | Adc34571 Human hep |
| 540 | 6 | 3.3 | 195 | 7 | ADM06147 | Adm06147 Human pro  | 613 | 6 | 3.3 | 208 | 7 | ADE56942 | Ade56942 Human pro |
| 541 | 6 | 3.3 | 196 | 3 | AAY95544 | Aay95544 Chlamydia  | 614 | 6 | 3.3 | 208 | 7 | ADE56946 | Ade56946 Human pro |
| 542 | 6 | 3.3 | 196 | 5 | ABP39441 | Abp39441 Staphyloc  | 615 | 6 | 3.3 | 208 | 7 | ADH91995 | Adh91995 Human hep |
| 543 | 6 | 3.3 | 196 | 8 | ADL06031 | Adl06031 M. catarr  | 616 | 6 | 3.3 | 208 | 7 | ADE29191 | Ade29191 Human hep |
| 544 | 6 | 3.3 | 197 | 2 | AAR43887 | Aar43887 Human lip  | 617 | 6 | 3.3 | 208 | 8 | ADF42792 | Adf42792 Human EGF |
| 545 | 6 | 3.3 | 197 | 2 | AAR68929 | Aar68929 Lipopolys  | 618 | 6 | 3.3 | 208 | 8 | ADN04000 | Adn04000 Antipsoi  |
| 546 | 6 | 3.3 | 197 | 2 | AAR68914 | Aar68914 Lipopolys  | 619 | 6 | 3.3 | 208 | 8 | ABB71753 | Abb71753 Drosophil |
| 547 | 6 | 3.3 | 197 | 2 | AAR68915 | Aar68915 Lipopolys  | 620 | 6 | 3.3 | 209 | 4 | ADA33868 | Ada33868 Acinetoba |
| 548 | 6 | 3.3 | 197 | 2 | AAR68912 | Aar68912 Human LBP  | 621 | 6 | 3.3 | 209 | 7 | ABO63954 | AbO63954 Klebsiell |
| 549 | 6 | 3.3 | 197 | 2 | AAR76602 | Aar76602 Recombina  | 622 | 6 | 3.3 | 210 | 6 | ABO63955 | AbO63955 Protein e |
| 550 | 6 | 3.3 | 197 | 2 | AAR67999 | Aar67999 rLBP25 N-  | 623 | 6 | 3.3 | 215 | 6 | ADA54191 | Ada54191 Human pro |
| 551 | 6 | 3.3 | 197 | 2 | AAW40164 | Aaw40164 LBP/BPI h  | 624 | 6 | 3.3 | 215 | 7 | ADJ57601 | Adj57601 Human apo |
| 552 | 6 | 3.3 | 197 | 2 | AAW40138 | Aaw40138 Human par  | 625 | 6 | 3.3 | 216 | 4 | AAW79734 | Aaw79734 Human pro |
| 553 | 6 | 3.3 | 197 | 2 | AAW40158 | Aaw40158 LBP/BPI h  | 626 | 6 | 3.3 | 216 | 6 | ABP75685 | Abp75685 Human sec |
| 554 | 6 | 3.3 | 197 | 5 | ABP73906 | Abp73906 Candida a  | 627 | 6 | 3.3 | 218 | 3 | AAW49865 | Aaw49865 Arabidops |
| 555 | 6 | 3.3 | 197 | 5 | AAE24828 | Aae24828 Human rLB  | 628 | 6 | 3.3 | 219 | 2 | AAV37699 | Aav37699 Chlamydia |
| 556 | 6 | 3.3 | 197 | 5 | AAE24808 | Aae24808 Human rLB  | 629 | 6 | 3.3 | 219 | 4 | AAU62595 | Aau62595 Propionib |
| 557 | 6 | 3.3 | 197 | 5 | AAE24834 | Aae24834 LBP deriv  | 630 | 6 | 3.3 | 219 | 5 | ABP41263 | Abp41263 Human ova |
| 558 | 6 | 3.3 | 197 | 5 | ABP57759 | Abp57759 E. coli L  | 631 | 6 | 3.3 | 219 | 6 | ABM59114 | Abm59114 Propionib |
| 559 | 6 | 3.3 | 197 | 6 | ABU28360 | Abu28360 Protein e  | 632 | 6 | 3.3 | 219 | 7 | ADB64770 | Adb64770 Human pro |
| 560 | 6 | 3.3 | 197 | 6 | ABU14966 | Abu14966 Protein e  | 633 | 6 | 3.3 | 220 | 3 | AAU50673 | Aau50673 Arabidops |
| 561 | 6 | 3.3 | 197 | 7 | ABW02250 | Abw02250 Human rLB  | 634 | 6 | 3.3 | 220 | 4 | ADM12952 | Adm12952 PRO83476  |
| 562 | 6 | 3.3 | 197 | 7 | ABW02229 | Abw02229 Human rLB  | 635 | 6 | 3.3 | 220 | 8 | ADM12951 | Adm12951 Antipsoi  |
| 563 | 6 | 3.3 | 197 | 7 | ABW02256 | Abw02256 LBP (lipo  | 636 | 6 | 3.3 | 221 | 3 | AAV32279 | Aav32279 Cat CD28  |
| 564 | 6 | 3.3 | 198 | 2 | AAR68930 | Aar68930 LBP (1-146 | 637 | 6 | 3.3 | 221 | 3 | AAV32279 | Aav32279 Cat CD28  |
| 565 | 6 | 3.3 | 198 | 2 | AAW40166 | Aaw40166 LBP/BPI h  | 638 | 6 | 3.3 | 221 | 3 | AAV32279 | Aav32279 Cat CD28  |
| 566 | 6 | 3.3 | 198 | 2 | AAW40165 | Aaw40165 LBP/BPI h  | 639 | 6 | 3.3 | 221 | 3 | AAV32286 | Aav32286 Feline CD |
| 567 | 6 | 3.3 | 198 | 5 | AAE24835 | Aae24835 LBP deriv  | 640 | 6 | 3.3 | 221 | 5 | AAO17735 | Aao17735 Feline CD |
| 568 | 6 | 3.3 | 198 | 5 | AAE24836 | Aae24836 LBP deriv  | 641 | 6 | 3.3 | 221 | 5 | AAU78122 | Aau78122 Feline CD |
| 569 | 6 | 3.3 | 198 | 7 | ABW02257 | Abw02257 LBP (lipo  | 642 | 6 | 3.3 | 221 | 7 | ADC01036 | Adc01036 Enterohae |
| 570 | 6 | 3.3 | 198 | 7 | ABW02258 | Abw02258 LBP (lipo  | 643 | 6 | 3.3 | 221 | 7 | ADC00895 | Adc00895 Enterohae |
| 571 | 6 | 3.3 | 199 | 2 | AAR68931 | Aar68931 LBP (1-87) | 644 | 6 | 3.3 | 222 | 6 | ABP79131 | Abp79131 N. gonorr |
| 572 | 6 | 3.3 | 199 | 2 | AAR68926 | Aar68926 Lipopolys  | 645 | 6 | 3.3 | 222 | 6 | ABP79131 | Abp79131 N. gonorr |
| 573 | 6 | 3.3 | 199 | 2 | AAR68926 | Aar68926 Lipopolys  | 646 | 6 | 3.3 | 224 | 6 | ABU27378 | Abu27378 Protein e |
| 574 | 6 | 3.3 | 199 | 2 | AAW40159 | Aaw40159 LBP/BPI h  | 647 | 6 | 3.3 | 224 | 8 | ABO60076 | AbO60076 Human gen |
| 575 | 6 | 3.3 | 199 | 2 | AAW40161 | Aaw40161 LBP/BPI h  | 648 | 6 | 3.3 | 226 | 4 | ABG15835 | Abg15835 Novel hum |
| 576 | 6 | 3.3 | 199 | 2 | AAV35177 | Aav35177 Chlamydia  | 649 | 6 | 3.3 | 226 | 6 | ABR56734 | AbR56734 Human sec |
| 577 | 6 | 3.3 | 199 | 4 | ABG68260 | Abg68260 Drosophil  | 650 | 6 | 3.3 | 226 | 8 | ABO60372 | AbO60372 Human gen |
| 578 | 6 | 3.3 | 199 | 5 | AAE24831 | Aae24831 Human LBP  | 651 | 6 | 3.3 | 226 | 8 | ABO58427 | AbO58427 Human gen |
| 579 | 6 | 3.3 | 199 | 5 | AAE24829 | Aae24829 Human LBP  | 652 | 6 | 3.3 | 229 | 3 | AAV59413 | Aav59413 Mouse CIP |
| 580 | 6 | 3.3 | 199 | 7 | ABW02253 | Abw02253 LBP (lipo  | 653 | 6 | 3.3 | 231 | 2 | AAV37511 | Aav37511 Protein w |
| 581 | 6 | 3.3 | 199 | 7 | ABW02251 | Abw02251 LBP (lipo  | 654 | 6 | 3.3 | 232 | 4 | AAU67716 | Aau67716 Propionib |
| 582 | 6 | 3.3 | 201 | 2 | AAO6839  | Aao6839 Human IL-   | 655 | 6 | 3.3 | 232 | 4 | AAU67716 | Aau67716 Propionib |
| 583 | 6 | 3.3 | 201 | 2 | AAV35941 | Aav35941 Extended   | 656 | 6 | 3.3 | 232 | 5 | ABU05964 | Abu05964 M. tuberc |
| 584 | 6 | 3.3 | 201 | 3 | AAU48961 | Aau48961 Arabidops  | 657 | 6 | 3.3 | 232 | 5 | ABU05964 | Abu05964 M. tuberc |
| 585 | 6 | 3.3 | 201 | 8 | ADP19249 | Adp19249 Human sec  | 658 | 6 | 3.3 | 233 | 3 | AAV59413 | Aav59413 Mouse CIP |
| 586 | 6 | 3.3 | 202 | 4 | AAU27496 | Aau27496 Human G-P  | 659 | 6 | 3.3 | 233 | 4 | ABG68857 | Abg68857 Drosophil |
| 587 | 6 | 3.3 | 204 | 3 | AAU76367 | Aau76367 Fragment   | 660 | 6 | 3.3 | 233 | 4 | ABG68857 | Abg68857 Drosophil |
| 588 | 6 | 3.3 | 204 | 4 | AAU42614 | Aau42614 Propionib  | 661 | 6 | 3.3 | 233 | 5 | ABP01339 | Abp01339 Putative  |
| 589 | 6 | 3.3 | 204 | 6 | ABM39133 | Abm39133 Propionib  | 662 | 6 | 3.3 | 233 | 5 | ABP01339 | Abp01339 Putative  |
| 590 | 6 | 3.3 | 204 | 7 | ADE12015 | Adel2015 Human sec  | 663 | 6 | 3.3 | 233 | 5 | ABP01339 | Abp01339 Putative  |
| 591 | 6 | 3.3 | 204 | 8 | ADO26507 | Ado26507 Ubiqutin   | 664 | 6 | 3.3 | 233 | 5 | ABP01339 | Abp01339 Putative  |
| 592 | 6 | 3.3 | 205 | 2 | AAR89338 | Aar89338 Minor tip  | 665 | 6 | 3.3 | 233 | 8 | AAU97933 | Aau97933 T. terres |
| 593 | 6 | 3.3 | 205 | 5 | ABG55248 | Abg55248 Lactococ   | 666 | 6 | 3.3 | 234 | 4 | AAU97933 | Aau97933 T. terres |
| 594 | 6 | 3.3 | 207 | 4 | AAU40198 | Aau40198 Propionib  | 667 | 6 | 3.3 | 234 | 5 | ABG09702 | Abg09702 S. epide  |
| 595 | 6 | 3.3 | 207 | 6 | ABM36717 | Abm36717 Propionib  | 668 | 6 | 3.3 | 235 | 7 | ADE63997 | Ade63997 Rat Prote |
| 596 | 6 | 3.3 | 207 | 7 | ADP03949 | Adp03949 Bacterial  | 669 | 6 | 3.3 | 235 | 7 | ADE63997 | Ade63997 Rat Prote |
| 597 | 6 | 3.3 | 207 | 7 | ADM27027 | Adm27027 Hyperther  | 670 | 6 | 3.3 | 236 | 1 | AAU91888 | Aau91888 Sequence  |
| 598 | 6 | 3.3 | 208 | 2 | AAR23998 | Aar23998 EGF/HB-EH  | 671 | 6 | 3.3 | 236 | 3 | AAV54331 | Aav54331 Amino aci |
| 599 | 6 | 3.3 | 208 | 2 | AAR80785 | Aar80785 Human pre  | 672 | 6 | 3.3 | 236 | 3 | AAV54331 | Aav54331 Amino aci |
| 600 | 6 | 3.3 | 208 | 3 | AAY58959 | Aay58959 Human HBE  | 673 | 6 | 3.3 | 236 | 6 | ABR41359 | AbR41359 Human DIT |
| 601 | 6 | 3.3 | 208 | 3 | AAU90406 | Aau90406 Human hep  | 674 | 6 | 3.3 | 237 | 6 | ABU38430 | Abu38430 Protein e |
| 602 | 6 | 3.3 | 208 | 3 | AAU90406 | Aau90406 Human HBE  | 675 | 6 | 3.3 | 237 | 6 | ABU38430 | Abu38430 Protein e |
| 603 | 6 | 3.3 | 208 | 4 | AAE15908 | Aae15908 Human HB-  | 676 | 6 | 3.3 | 238 | 4 | ABG59481 | Abg59481 Drosophil |
| 604 | 6 | 3.3 | 208 | 5 | AAE15908 | Aae15908 Human HB-  | 677 | 6 | 3.3 | 238 | 5 | AAO14223 | Aao14223 Drosophil |
| 605 | 6 | 3.3 | 208 | 5 | AAE25775 | Aae25775 Human dip  | 678 | 6 | 3.3 | 238 | 6 | ABU19767 | Abu19767 Protein e |
| 606 | 6 | 3.3 | 208 | 6 | ABP96795 | Abp96795 Human COP  | 679 | 6 | 3.3 | 238 | 6 | ABU22091 | Abu22091 Protein e |
| 607 | 6 | 3.3 | 208 | 6 | ABM66997 | Abm66997 Phototrab  | 680 | 6 | 3.3 | 238 | 6 | ABU22091 | Abu22091 Protein e |
| 608 | 6 | 3.3 | 208 | 6 | ABU07422 | Abu07422 Protein d  | 681 | 6 | 3.3 | 238 | 6 | ADA35364 | Ada35364 Acinetoba |
| 609 | 6 | 3.3 | 208 | 6 | ABU09519 | Abu09519 Human Hep  | 682 | 6 | 3.3 | 240 | 4 | AAU72699 | Aau72699 Human bet |

|     |   |     |     |   |          |                     |     |   |     |     |   |           |                     |
|-----|---|-----|-----|---|----------|---------------------|-----|---|-----|-----|---|-----------|---------------------|
| 683 | 6 | 3.3 | 240 | 7 | ADC31558 | ADC31558 Human nov  | 756 | 6 | 3.3 | 273 | 4 | AAB80358  | Aab80358 Human pro  |
| 684 | 6 | 3.3 | 240 | 7 | ADM26671 | Adm26671 Hyperther  | 757 | 6 | 3.3 | 273 | 4 | AAG81495  | Aag81495 S. epider  |
| 685 | 6 | 3.3 | 241 | 3 | AG06531  | Ag06531 Arabidops   | 758 | 6 | 3.3 | 273 | 5 | ABP28059  | Abp28059 Streptoco  |
| 686 | 6 | 3.3 | 241 | 6 | ABP78972 | Abp78972 N. gonorr  | 759 | 6 | 3.3 | 273 | 7 | ABO67947  | Ab067947 Pseudomon  |
| 687 | 6 | 3.3 | 242 | 5 | ABP45819 | Abp45819 Human Bly  | 760 | 6 | 3.3 | 273 | 8 | ADN46606  | Adn46606 Thermococ  |
| 688 | 6 | 3.3 | 242 | 7 | ADG96646 | Adg96646 Single ch  | 761 | 6 | 3.3 | 274 | 2 | AAI06813  | Aai06813 R. pseudo  |
| 689 | 6 | 3.3 | 242 | 8 | ADN05213 | Adn05213 Antipsori  | 762 | 6 | 3.3 | 275 | 5 | ABP38887  | Abp38887 Staphyloc  |
| 690 | 6 | 3.3 | 242 | 8 | ABO59348 | Ab059348 Human gen  | 763 | 6 | 3.3 | 275 | 5 | ABU05951  | Abu05951 M. tuberc  |
| 691 | 6 | 3.3 | 242 | 8 | ABO59348 | Ab059348 Human gen  | 764 | 6 | 3.3 | 275 | 5 | ABP77986  | Abp77986 N. gonorr  |
| 692 | 6 | 3.3 | 244 | 2 | AAV05823 | Aav05823 Arabidops  | 765 | 6 | 3.3 | 276 | 7 | ADC87461  | Adc87461 Human GPC  |
| 693 | 6 | 3.3 | 244 | 4 | ABB68407 | Abb68407 Drosophil  | 766 | 6 | 3.3 | 276 | 7 | ABG00520  | Abg00520 Novel hum  |
| 694 | 6 | 3.3 | 244 | 6 | ABB82785 | Abb82785 Human 3OS  | 767 | 6 | 3.3 | 277 | 4 | ABU49447  | Abu49447 Protein e  |
| 695 | 6 | 3.3 | 245 | 5 | AAE20289 | Aae20289 Chlamydia  | 768 | 6 | 3.3 | 278 | 4 | ABG26687  | Abg26687 Novel hum  |
| 696 | 6 | 3.3 | 246 | 4 | ABB12096 | Abb12096 Human APE  | 769 | 6 | 3.3 | 280 | 3 | AAG06529  | Aag06529 Arabidops  |
| 697 | 6 | 3.3 | 246 | 6 | ABU40562 | Abu40562 Protein e  | 770 | 6 | 3.3 | 281 | 3 | AAG06529  | Aag06529 Arabidops  |
| 698 | 6 | 3.3 | 247 | 7 | ADP60194 | Adp60194 Human con  | 771 | 6 | 3.3 | 281 | 3 | AAB58317  | Aab58317 Lung canc  |
| 699 | 6 | 3.3 | 247 | 7 | ADH87337 | Adh87337 Enterococ  | 772 | 6 | 3.3 | 282 | 8 | ADN73633  | Adn73633 Thale cre  |
| 700 | 6 | 3.3 | 249 | 3 | AAG28458 | Aag28458 Arabidops  | 773 | 6 | 3.3 | 283 | 6 | ADN73633  | Adn73633 Thale cre  |
| 701 | 6 | 3.3 | 250 | 4 | AAU00212 | Aau00212 Monoclonia | 774 | 6 | 3.3 | 284 | 6 | ABU34240  | Abu34240 Human can  |
| 702 | 6 | 3.3 | 250 | 4 | AAG91194 | Ag91194 C glutami   | 775 | 6 | 3.3 | 284 | 6 | ABU34240  | Abu34240 Human can  |
| 703 | 6 | 3.3 | 250 | 4 | ABG10352 | Abg10352 Novel hum  | 776 | 6 | 3.3 | 285 | 7 | ABR62847  | AbR62847 Human car  |
| 704 | 6 | 3.3 | 250 | 6 | ABU22377 | Abu22377 Protein e  | 777 | 6 | 3.3 | 285 | 7 | ADE08420  | Ade08420 Novel pro  |
| 705 | 6 | 3.3 | 252 | 2 | AAI35557 | Aai35557 Thermococ  | 778 | 6 | 3.3 | 286 | 4 | ABB65966  | Abb65966 Drosophil  |
| 706 | 6 | 3.3 | 252 | 6 | ABJ38336 | Abj38336 TALL-1 in  | 779 | 6 | 3.3 | 286 | 4 | ABG81654  | Abg81654 S. epider  |
| 707 | 6 | 3.3 | 252 | 6 | ABM66089 | Abm66089 Propionib  | 780 | 6 | 3.3 | 288 | 5 | ABB55106  | Abb55106 Lactococc  |
| 708 | 6 | 3.3 | 252 | 7 | ABR39192 | AbR39192 Chimeric   | 781 | 6 | 3.3 | 288 | 6 | ABM72654  | Abm72654 Staphyloc  |
| 709 | 6 | 3.3 | 253 | 3 | AG42592  | Ag42592 Arabidops   | 782 | 6 | 3.3 | 288 | 8 | ADI39298  | Adi39298 S. hygro   |
| 710 | 6 | 3.3 | 253 | 3 | AG06530  | Ag06530 Arabidops   | 783 | 6 | 3.3 | 288 | 8 | ADI39266  | Adi39266 S. hygro   |
| 711 | 6 | 3.3 | 254 | 7 | ADF04806 | Adf04806 Bacterial  | 784 | 6 | 3.3 | 289 | 6 | ABU44122  | Abu44122 Protein e  |
| 712 | 6 | 3.3 | 254 | 8 | ADM41480 | Adm41480 Perennial  | 785 | 6 | 3.3 | 290 | 5 | ABB48735  | Abb48735 Listeria   |
| 713 | 6 | 3.3 | 255 | 5 | ABE50068 | AbE50068 Listeria   | 786 | 6 | 3.3 | 290 | 7 | ADE15974  | Ade15974 G-coupled  |
| 714 | 6 | 3.3 | 256 | 7 | ABO23517 | Ab023517 Mycobacte  | 787 | 6 | 3.3 | 290 | 7 | ABO72971  | Ab072971 Pseudomon  |
| 715 | 6 | 3.3 | 256 | 8 | ADM41482 | Adm41482 Perennial  | 788 | 6 | 3.3 | 291 | 2 | ADL93913  | Adl93913 Human G-c  |
| 716 | 6 | 3.3 | 257 | 4 | AAU63320 | Aau63320 Propionib  | 789 | 6 | 3.3 | 291 | 8 | AAV17070  | Aav17070 C. elega   |
| 717 | 6 | 3.3 | 257 | 6 | ABM59839 | Abm59839 Propionib  | 790 | 6 | 3.3 | 291 | 3 | ABR32545  | AbR32545 Eucalyptu  |
| 718 | 6 | 3.3 | 257 | 7 | ADC31659 | Adc31659 Human nov  | 791 | 6 | 3.3 | 293 | 3 | ABR32545  | AbR32545 Eucalyptu  |
| 719 | 6 | 3.3 | 258 | 7 | ABE12021 | AbE12021 Human sec  | 792 | 6 | 3.3 | 293 | 3 | ABR32545  | AbR32545 Eucalyptu  |
| 720 | 6 | 3.3 | 260 | 4 | ABG13338 | Abg13338 Novel hum  | 793 | 6 | 3.3 | 293 | 3 | ABR32545  | AbR32545 Eucalyptu  |
| 721 | 6 | 3.3 | 260 | 7 | ABO72642 | Ab072642 Pseudomon  | 794 | 6 | 3.3 | 295 | 7 | ADL27199  | Adl27199 Amino aci  |
| 722 | 6 | 3.3 | 262 | 5 | ABB54686 | Abb54686 Lactococc  | 795 | 6 | 3.3 | 296 | 5 | ABP39628  | Abp39628 Staphyloc  |
| 723 | 6 | 3.3 | 262 | 6 | ADA35265 | Ada35265 Acinetoba  | 796 | 6 | 3.3 | 296 | 5 | ABP39628  | Abp39628 Staphyloc  |
| 724 | 6 | 3.3 | 262 | 7 | ADC30986 | Adc30986 Human nov  | 797 | 6 | 3.3 | 296 | 8 | ABM71979  | Abm71979 Staphyloc  |
| 725 | 6 | 3.3 | 263 | 8 | ADM41481 | Adm41481 Tall fesc  | 798 | 6 | 3.3 | 297 | 4 | AAAB63995 | AaAB63995 Drosophil |
| 726 | 6 | 3.3 | 263 | 6 | ABU57634 | Abu57634 Differen   | 799 | 6 | 3.3 | 297 | 4 | AAAB63995 | AaAB63995 Drosophil |
| 727 | 6 | 3.3 | 264 | 4 | ABB58972 | Abb58972 Drosophil  | 800 | 6 | 3.3 | 297 | 4 | AAAB63995 | AaAB63995 Drosophil |
| 728 | 6 | 3.3 | 264 | 7 | ABO81346 | Ab081346 Pseudomon  | 801 | 6 | 3.3 | 299 | 2 | AAW59130  | Aaw59130 Homo sapi  |
| 729 | 6 | 3.3 | 265 | 2 | ABR29942 | AbR29942 Deduced f  | 802 | 6 | 3.3 | 300 | 4 | ABR12019  | AbR12019 Human pol  |
| 730 | 6 | 3.3 | 265 | 2 | AAK94723 | Aak94723 PRRSV Iel  | 803 | 6 | 3.3 | 301 | 4 | ABR12019  | AbR12019 Human pol  |
| 731 | 6 | 3.3 | 265 | 5 | ABP30338 | Abp30338 Streptoco  | 804 | 6 | 3.3 | 301 | 6 | ABU40262  | Abu40262 Protein e  |
| 732 | 6 | 3.3 | 265 | 5 | ABR38823 | AbR38823 A. niger   | 805 | 6 | 3.3 | 301 | 6 | ADA36490  | Ada36490 Acinetoba  |
| 733 | 6 | 3.3 | 265 | 5 | ABR38823 | AbR38823 A. niger   | 806 | 6 | 3.3 | 301 | 7 | AAO30411  | Aao30411 Human sec  |
| 734 | 6 | 3.3 | 265 | 6 | ABP56118 | Abp56118 Bothrops   | 807 | 6 | 3.3 | 301 | 7 | AAO30411  | Aao30411 Human sec  |
| 735 | 6 | 3.3 | 266 | 7 | ADH88206 | Adh88206 Enterococ  | 808 | 6 | 3.3 | 302 | 3 | AAV75484  | Aav75484 Neisseria  |
| 736 | 6 | 3.3 | 267 | 3 | AAO08391 | Aao08391 Arabidops  | 809 | 6 | 3.3 | 302 | 3 | AAV75484  | Aav75484 Neisseria  |
| 737 | 6 | 3.3 | 267 | 4 | AAU35956 | Aau35956 Helicobac  | 810 | 6 | 3.3 | 302 | 7 | ADH87331  | Adh87331 Enterococ  |
| 738 | 6 | 3.3 | 267 | 7 | ABO83901 | Ab083901 Pseudomon  | 811 | 6 | 3.3 | 303 | 4 | ADM19794  | Adm19794 Protein e  |
| 739 | 6 | 3.3 | 268 | 5 | ABB48434 | Abb48434 Listeria   | 812 | 6 | 3.3 | 303 | 6 | ABU23165  | Abu23165 Protein e  |
| 740 | 6 | 3.3 | 268 | 6 | ABU02698 | Abu02698 S. pneumo  | 813 | 6 | 3.3 | 304 | 4 | ABR11508  | AbR11508 Human C8   |
| 741 | 6 | 3.3 | 268 | 6 | ABP81496 | Abp81496 Streptoco  | 814 | 6 | 3.3 | 304 | 4 | ABG16574  | Abg16574 Novel hum  |
| 742 | 6 | 3.3 | 268 | 8 | ADK47789 | Adk47789 Streptoco  | 815 | 6 | 3.3 | 304 | 5 | ABP26878  | Abp26878 Streptoco  |
| 743 | 6 | 3.3 | 268 | 8 | ADM92267 | Adm92267 S pneumon  | 816 | 6 | 3.3 | 305 | 4 | ABG04250  | Abg04250 Novel hum  |
| 744 | 6 | 3.3 | 269 | 4 | ADG74550 | Adg74550 Human col  | 817 | 6 | 3.3 | 305 | 5 | ABH80780  | Abh80780 Rat stem   |
| 745 | 6 | 3.3 | 269 | 8 | ADJ92707 | Adj92707 Corn beta  | 818 | 6 | 3.3 | 305 | 7 | ADE62974  | Ade62974 Rat Prote  |
| 746 | 6 | 3.3 | 269 | 8 | ADQ92312 | Adq92312 Maize bet  | 819 | 6 | 3.3 | 305 | 7 | ADP17716  | Adp17716 Rat UDP-g  |
| 747 | 6 | 3.3 | 270 | 2 | AAW20644 | Aaw20644 H. pylori  | 820 | 6 | 3.3 | 305 | 7 | ABO68956  | AbO68956 Pseudomon  |
| 748 | 6 | 3.3 | 270 | 4 | ABR61318 | AbR61318 Drosophil  | 821 | 6 | 3.3 | 306 | 7 | ABO68956  | AbO68956 Pseudomon  |
| 749 | 6 | 3.3 | 270 | 5 | ADP63093 | Adp63093 FLOII gen  | 822 | 6 | 3.3 | 306 | 7 | ABO69325  | AbO69325 Pseudomon  |
| 750 | 6 | 3.3 | 270 | 7 | ADC94526 | Adc94526 E. faeciu  | 823 | 6 | 3.3 | 308 | 4 | AAU04817  | Aau04817 Micromono  |
| 751 | 6 | 3.3 | 271 | 3 | ABR44523 | AbR44523 Virulence  | 824 | 6 | 3.3 | 308 | 5 | AAU04817  | Aau04817 Micromono  |
| 752 | 6 | 3.3 | 271 | 5 | ABP54475 | Abp54475 Pasteurel  | 825 | 6 | 3.3 | 308 | 6 | ADP99219  | Adp99219 Orthosomy  |
| 753 | 6 | 3.3 | 271 | 7 | ADC86627 | Adc86627 Human GPC  | 826 | 6 | 3.3 | 308 | 7 | ABO76070  | AbO76070 Pseudomon  |
| 754 | 6 | 3.3 | 271 | 8 | ADQ43394 | Adq43394 Lutromyia  | 827 | 6 | 3.3 | 309 | 6 | ABU39854  | Abu39854 Protein e  |
| 755 | 6 | 3.3 | 272 | 6 | ABM64917 | Abm64917 Propionib  | 828 | 6 | 3.3 | 310 | 4 | AAW93226  | Aaw93226 Human pol  |



|     |   |      |     |   |           |           |           |     |   |      |     |   |           |          |           |
|-----|---|------|-----|---|-----------|-----------|-----------|-----|---|------|-----|---|-----------|----------|-----------|
| 829 | 6 | 3..3 | 310 | 4 | AAB76840  | Aab76840  | Corynebac | 902 | 6 | 3..3 | 332 | 7 | ADF04360  | Adf04360 | Bacterial |
| 830 | 6 | 3..3 | 310 | 5 | AAU77137  | Aau77137  | Human alp | 903 | 6 | 3..3 | 332 | 8 | ADJ48487  | Adj48487 | Maize oil |
| 831 | 6 | 3..3 | 310 | 5 | ABP43541  | Abp43541  | Human sec | 904 | 6 | 3..3 | 332 | 8 | ADJ49812  | Adj49812 | Oil-ssoc  |
| 832 | 6 | 3..3 | 310 | 5 | ABR38856  | AbR38856  | A. niger  | 905 | 6 | 3..3 | 333 | 3 | AAy96275  | Aay96275 | Mammalian |
| 833 | 6 | 3..3 | 310 | 7 | ADL15976  | Adel15976 | G-coupled | 906 | 6 | 3..3 | 333 | 4 | AAG71417  | Aag71417 | Human oil |
| 834 | 6 | 3..3 | 310 | 7 | ADL65503  | Adl65503  | C. glutam | 907 | 6 | 3..3 | 333 | 4 | AAG71684  | Aag71684 | Human oil |
| 835 | 6 | 3..3 | 310 | 8 | ADL30608  | Adl30608  | Human pro | 908 | 6 | 3..3 | 333 | 4 | AAG71422  | Aag71422 | Human oil |
| 836 | 6 | 3..3 | 310 | 8 | ADL93915  | Adl93915  | Human G-c | 909 | 6 | 3..3 | 333 | 4 | AAG71712  | Aag71712 | Human oil |
| 837 | 6 | 3..3 | 312 | 4 | ABB64002  | Abb64002  | Drosophil | 910 | 6 | 3..3 | 333 | 5 | ABG34089  | Abg34089 | Human Tax |
| 838 | 6 | 3..3 | 312 | 4 | AAG71730  | Aag71730  | Human oil | 911 | 6 | 3..3 | 333 | 5 | ABG34127  | Abg34127 | Human Tax |
| 839 | 6 | 3..3 | 312 | 4 | AAG71996  | Aag71996  | Human oil | 912 | 6 | 3..3 | 333 | 7 | ADD45181  | Add45181 | Human GLU |
| 840 | 6 | 3..3 | 312 | 4 | AAG72264  | Aag72264  | Human oil | 913 | 6 | 3..3 | 333 | 7 | ADD45486  | Add45486 | Human Pro |
| 841 | 6 | 3..3 | 312 | 5 | ABG76821  | Abg76821  | Human G-p | 914 | 6 | 3..3 | 333 | 7 | ADE54860  | Ades4860 | Human Pro |
| 842 | 6 | 3..3 | 312 | 5 | ABG76822  | Abg76822  | Human G-p | 915 | 6 | 3..3 | 333 | 7 | ADE54864  | Ades4864 | Human Pro |
| 843 | 6 | 3..3 | 312 | 5 | ABP51592  | Abp51592  | Human G-p | 916 | 6 | 3..3 | 334 | 3 | ABE54868  | Abes4868 | Human Pro |
| 844 | 6 | 3..3 | 312 | 7 | ADL04309  | Adl04309  | Human G-p | 917 | 6 | 3..3 | 334 | 4 | AU20446   | Aau20446 | Human pro |
| 845 | 6 | 3..3 | 313 | 4 | ABG30006  | Abg30006  | Novel hum | 918 | 6 | 3..3 | 334 | 4 | AU18360   | Aau18360 | Human sec |
| 846 | 6 | 3..3 | 313 | 5 | ABR89780  | AbR89780  | Human pol | 919 | 6 | 3..3 | 334 | 7 | ADM05994  | Adm05994 | Human end |
| 847 | 6 | 3..3 | 314 | 7 | ABO63113  | AbO63113  | Klebsiell | 920 | 6 | 3..3 | 336 | 2 | AAW79094  | Aaw79094 | Human pro |
| 848 | 6 | 3..3 | 315 | 5 | ABP38716  | Abp38716  | Staphyloc | 921 | 6 | 3..3 | 336 | 5 | ABP61795  | Abp61795 | Human pol |
| 849 | 6 | 3..3 | 317 | 7 | ABO74030  | AbO74030  | Pseudomon | 922 | 6 | 3..3 | 336 | 6 | ADA56785  | Ada56785 | Human sec |
| 850 | 6 | 3..3 | 317 | 8 | ADP74135  | Adp74135  | A. fulgid | 923 | 6 | 3..3 | 336 | 6 | ABR47665  | AbR47665 | Human sec |
| 851 | 6 | 3..3 | 317 | 8 | ADP74133  | Adp74133  | A. fulgid | 924 | 6 | 3..3 | 336 | 6 | ABR00029  | AbR00029 | Human gen |
| 852 | 6 | 3..3 | 318 | 7 | ABM74281  | AbM74281  | DNA clone | 925 | 6 | 3..3 | 336 | 7 | ADB91478  | AdB91478 | Human sec |
| 853 | 6 | 3..3 | 319 | 6 | ABU23557  | AbU23557  | Protein e | 926 | 6 | 3..3 | 336 | 7 | ADC74043  | AdC74043 | Human sec |
| 854 | 6 | 3..3 | 320 | 4 | ADA32968  | Ada32968  | Acinetoba | 927 | 6 | 3..3 | 336 | 7 | ADD37841  | Add37841 | Human sec |
| 855 | 6 | 3..3 | 320 | 4 | ABR58987  | AbR58987  | Drosophil | 928 | 6 | 3..3 | 337 | 2 | ADY07754  | Aay07754 | Human sec |
| 856 | 6 | 3..3 | 322 | 4 | ABR60043  | AbR60043  | Drosophil | 929 | 6 | 3..3 | 337 | 6 | ABU40104  | AbU40104 | Protein e |
| 857 | 6 | 3..3 | 322 | 4 | ABR60976  | AbB60976  | Drosophil | 930 | 6 | 3..3 | 338 | 3 | AAy97541  | Aay97541 | Rat KIAA0 |
| 858 | 6 | 3..3 | 322 | 8 | ADO70022  | AdO70022  | Human ins | 931 | 6 | 3..3 | 340 | 4 | AAU35143  | Aau35143 | Enterococ |
| 859 | 6 | 3..3 | 323 | 5 | ABR60922  | AbP60922  | Sulfolobu | 932 | 6 | 3..3 | 341 | 4 | AAAB80160 | Aab80160 | Corynebac |
| 860 | 6 | 3..3 | 323 | 5 | ABR58655  | AbR58655  | Human can | 933 | 6 | 3..3 | 341 | 4 | AAAG92064 | Aag92064 | C glutam  |
| 861 | 6 | 3..3 | 324 | 7 | ABO64735  | AbO64735  | Klebsiell | 934 | 6 | 3..3 | 341 | 7 | ADD13487  | Add13487 | C. glutam |
| 862 | 6 | 3..3 | 325 | 6 | ABR82758  | AbB82758  | Human IGF | 935 | 6 | 3..3 | 341 | 7 | ADH86290  | AdH86290 | Enterococ |
| 863 | 6 | 3..3 | 325 | 7 | ADSR8187  | AdS8187   | Human Pro | 936 | 6 | 3..3 | 342 | 6 | ADA06231  | Ada06231 | Rice chol |
| 864 | 6 | 3..3 | 325 | 7 | ADSR8187  | AdS8187   | Human Pro | 937 | 6 | 3..3 | 342 | 8 | ADO17002  | AdO17002 | Rice chol |
| 865 | 6 | 3..3 | 325 | 7 | ADE54201  | AdE54201  | Human Pro | 938 | 6 | 3..3 | 344 | 2 | AAW06600  | Aaw06600 | Hyperseus |
| 866 | 6 | 3..3 | 326 | 4 | AAU34980  | Aau34980  | Enterococ | 939 | 6 | 3..3 | 344 | 2 | AAW61116  | Aaw61116 | Hyperseus |
| 867 | 6 | 3..3 | 326 | 6 | AAU29293  | Aau29293  | Protein e | 940 | 6 | 3..3 | 344 | 2 | AAW62457  | Aaw62457 | Hyperseus |
| 868 | 6 | 3..3 | 327 | 2 | AAy05529  | Aay05529  | Soybean v | 941 | 6 | 3..3 | 344 | 2 | AAW75865  | Aaw75865 | Pseudomon |
| 869 | 6 | 3..3 | 327 | 3 | AAy97832  | Aay97832  | Soybean v | 942 | 6 | 3..3 | 344 | 2 | AAW87841  | Aaw87841 | A hyperse |
| 870 | 6 | 3..3 | 327 | 4 | ABR66095  | AbB66095  | Drosophil | 943 | 6 | 3..3 | 344 | 3 | AAy71099  | Aay71099 | Pseudomon |
| 871 | 6 | 3..3 | 327 | 4 | ABR66095  | AbB66095  | Drosophil | 944 | 6 | 3..3 | 344 | 3 | AAy84860  | Aay84860 | A hyperse |
| 872 | 6 | 3..3 | 327 | 7 | ADJ32837  | Adj32837  | Soybean v | 945 | 6 | 3..3 | 344 | 3 | AAAB44563 | Aab44563 | Virulence |
| 873 | 6 | 3..3 | 328 | 2 | AAAR04908 | Aar04908  | EcoRI-Eco | 946 | 6 | 3..3 | 344 | 3 | AAAG27323 | Aag27323 | Arabidops |
| 874 | 6 | 3..3 | 328 | 2 | AAAR04908 | Aar04908  | EcoRI-Eco | 947 | 6 | 3..3 | 344 | 5 | AAO22551  | Aao22551 | Hyperseus |
| 875 | 6 | 3..3 | 328 | 4 | ABR59877  | AbR59877  | IGFBP-2 p | 948 | 6 | 3..3 | 344 | 5 | AAE18299  | Aae18299 | Pseudomon |
| 876 | 6 | 3..3 | 328 | 4 | AAAB49991 | Aab49991  | Human ins | 949 | 6 | 3..3 | 344 | 5 | ABP54515  | Abp54515 | Pasteurel |
| 877 | 6 | 3..3 | 328 | 5 | AAO17354  | Aao17354  | Human ins | 950 | 6 | 3..3 | 344 | 5 | AAE16453  | Aae16453 | P. solana |
| 878 | 6 | 3..3 | 328 | 6 | ABR6805   | AbP6805   | Human COP | 951 | 6 | 3..3 | 344 | 5 | ABB09229  | Abb09229 | P. solana |
| 879 | 6 | 3..3 | 328 | 6 | ABR82755  | AbB82755  | Human IGF | 952 | 6 | 3..3 | 344 | 6 | ABU39229  | Abu39229 | Protein e |
| 880 | 6 | 3..3 | 328 | 6 | ABR56585  | AbU56585  | Lung can  | 953 | 6 | 3..3 | 344 | 6 | ADAO6235  | Ada06235 | Wheat cho |
| 881 | 6 | 3..3 | 328 | 7 | ADJ14079  | Adj14079  | Human SRC | 954 | 6 | 3..3 | 344 | 6 | ADAO6235  | Ada06235 | Wheat cho |
| 882 | 6 | 3..3 | 328 | 7 | ADSR6686  | Ades6686  | Human Pro | 955 | 6 | 3..3 | 344 | 6 | ADC87149  | AdC87149 | Human GPC |
| 883 | 6 | 3..3 | 328 | 7 | ADSR6686  | Ades6686  | Human Pro | 956 | 6 | 3..3 | 344 | 8 | ADJ48705  | Adj48705 | Oil-assoc |
| 884 | 6 | 3..3 | 328 | 7 | ADJ68290  | Adj68290  | Human hea | 957 | 6 | 3..3 | 344 | 8 | ADO17006  | AdO17006 | Wheat cho |
| 885 | 6 | 3..3 | 328 | 8 | ADN95269  | AdN95269  | Human BPC | 958 | 6 | 3..3 | 349 | 6 | ADA06215  | Ada06215 | Corn chol |
| 886 | 6 | 3..3 | 328 | 8 | ADK70455  | AdK70455  | Respirato | 959 | 6 | 3..3 | 349 | 6 | ADA06225  | Ada06225 | Corn chol |
| 887 | 6 | 3..3 | 328 | 8 | ADL61308  | AdL61308  | Human ins | 960 | 6 | 3..3 | 349 | 8 | ADO16986  | AdO16986 | Corn chol |
| 888 | 6 | 3..3 | 328 | 8 | ADQ20663  | AdQ20663  | Human sof | 961 | 6 | 3..3 | 349 | 8 | ADO16996  | AdO16996 | Corn chol |
| 889 | 6 | 3..3 | 328 | 8 | ADQ26071  | AdQ26071  | Insulin-1 | 962 | 6 | 3..3 | 351 | 2 | AAW92796  | Aaw92796 | Tobacco T |
| 890 | 6 | 3..3 | 329 | 5 | AAU98469  | Aau98469  | Transcrip | 963 | 6 | 3..3 | 351 | 2 | ABP71036  | Abp71036 | Tobacco q |
| 891 | 6 | 3..3 | 329 | 6 | ADA06237  | Ada06237  | Rape chol | 964 | 6 | 3..3 | 351 | 7 | ABO76351  | AbO76351 | Pseudomon |
| 892 | 6 | 3..3 | 329 | 8 | ADJ49169  | Adj49169  | Oil-assoc | 965 | 6 | 3..3 | 352 | 3 | AAAG37965 | Aag37965 | Arabidops |
| 893 | 6 | 3..3 | 329 | 8 | ADK71002  | AdK71002  | Human est | 966 | 6 | 3..3 | 352 | 3 | AAAG09558 | Aag09558 | Arabidops |
| 894 | 6 | 3..3 | 329 | 8 | ADOL17008 | AdOl17008 | Brassica  | 967 | 6 | 3..3 | 352 | 3 | ADA54732  | Ada54732 | Human pro |
| 895 | 6 | 3..3 | 331 | 5 | ABR47655  | AbB47655  | Listeria  | 968 | 6 | 3..3 | 353 | 8 | ADH13678  | AdH13678 | Human ENZ |
| 896 | 6 | 3..3 | 331 | 6 | ABU32452  | Abu32452  | Protein e | 969 | 6 | 3..3 | 353 | 8 | ADK69718  | AdK69718 | Rhizopus  |
| 897 | 6 | 3..3 | 331 | 8 | ADA06236  | Ada06236  | Rape chol | 970 | 6 | 3..3 | 355 | 3 | AAy75279  | Aay75279 | Neisseria |
| 898 | 6 | 3..3 | 331 | 8 | ADJ49170  | Adj49170  | Oil-assoc | 971 | 6 | 3..3 | 355 | 3 | AAy75278  | Aay75278 | Neisseria |
| 899 | 6 | 3..3 | 331 | 8 | ADOL17007 | AdOl17007 | Brassica  | 972 | 6 | 3..3 | 355 | 3 | AAAG21670 | Aag21670 | Arabidops |
| 900 | 6 | 3..3 | 332 | 3 | AAAG06757 | Aag06757  | Arabidops | 973 | 6 | 3..3 | 355 | 6 | ADA09891  | Ada09891 | Human rec |
| 901 | 6 | 3..3 | 332 | 3 | AAAG32600 | Aag32600  | Arabidops | 974 | 6 | 3..3 | 355 | 7 | ABO72256  | AbO72256 | Pseudomon |

975 6 3.3 355 7 ABO82879 Pseudomon  
976 6 3.3 356 8 ABR47489 Breast ca  
977 6 3.3 356 8 ADL35513 Human hom  
978 6 3.3 357 8 ABU33221 Protein e  
979 6 3.3 357 8 ADK69725 Rhizopus  
980 6 3.3 357 8 ADK47483 Streptoco  
981 6 3.3 359 4 AAB80054 Corynebact  
982 6 3.3 359 4 AAG81175 Mycobacte  
983 6 3.3 359 4 AAG90184 C glutami  
984 6 3.3 359 6 ABU35861 Protein e  
985 6 3.3 359 6 ABU36730 Protein e  
986 6 3.3 359 6 AAO31002 Human tra  
987 6 3.3 359 6 AAO31018 Human tra  
988 6 3.3 359 7 AAD13509 C. glutam  
989 6 3.3 360 3 ABP40989 Human ret  
990 6 3.3 360 3 AAG05532 Arabidops  
991 6 3.3 360 8 ADL00257 Carotene-  
992 6 3.3 361 6 ADA35288 Acinetoba  
993 6 3.3 361 8 ADO20182 Human PRO  
994 6 3.3 362 4 AAY97745 Human CHR  
995 6 3.3 362 8 ADG22632 Cyanophag  
996 6 3.3 362 8 ADM90921 Human pha  
997 6 3.3 364 3 AAG09557 Arabidops  
998 6 3.3 364 3 AAG37964 Arabidops  
999 6 3.3 364 3 AAG21669 Arabidops  
1000 6 3.3 364 4 AAB94093 Human pro

## ALIGNMENTS

RESULT 1  
ID AAB47008 standard; protein; 180 AA.  
XX AAB47008;  
XX 22-MAR-2001 (first entry)  
XX L. intracellularis SodC.  
XX Porcine proliferative enteropathy; immunogen; SodC; antibody; pig;  
XX vaccine; intestinal infection; serum; blood lymph node; ileum; caecum;  
XX small intestine; large intestine; faeces; rectal swab; PPE.  
XX Lawsonia intracellularis.  
XX Key Location/Qualifiers  
XX Peptide 1..42  
XX /note= "Immunogenic peptide fragment"  
XX WO200069903-A1.  
XX 23-NOV-2000.  
XX 11-MAY-2000; 2000WO-AU000436.  
XX 13-MAY-1999; 99US-0133989P.  
XX (PFIZ ) PFIZER PROD INC.  
XX (FIGR-) FIG RES & DEV CORP.  
XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
XX Ankenbauer RG, Hasse D, Panaccio M, Rosey EL, Wright C;  
XX WPI; 2001-031924/04.  
XX N-ESDB; AAC85254.

Isolated or recombinant polypeptide for treating porcine and avian  
PT species against Lawsonia intracellularis infection, comprises, mimics or  
PT cross-reacts with the B or T cell epitope of Lawsonia SodC polypeptide.  
XX  
XX Claim 6; Page 79-80; 85pp; English.

XX This sequence represents an immunogenic polypeptide, SodC, which is  
CC capable of eliciting the production of antibodies against L.  
CC intracellularis when administered to an avian or porcine animal. This  
CC polypeptide can be used in a vaccine composition for the prophylaxis or  
CC treatment of intestinal infection of an animal by Lawsonia. The DNA  
CC encoding SodC polypeptide may be used in a vaccine vector such that the  
CC SodC polypeptide is expressible at a level sufficient to confer immunity  
CC against Lawsonia. The polypeptide is useful for diagnosing infection of a  
CC porcine or avian animal or identifying whether or not the animal has  
CC suffered from a past infection or is currently infected with L.  
CC intracellularis or a microorganism that is immunologically cross-reactive  
CC to it. This is done by contacting whole serum, blood lymph nodes, ileum,  
CC caecum, small intestine, large intestine, faeces or rectal swab derived  
CC from the animal with the immunogenic polypeptide or an antibody raised  
CC against it. Probes derived from the Lawsonia DNA are useful for detecting  
CC L. intracellularis or related microorganisms by hybridization. The SodC  
CC polypeptide is useful in the preparation of a medicament for the  
CC treatment and prophylaxis of porcine proliferative enteropathy (PPE)  
XX  
XX SQ Sequence 180 AA;

Query Match 100.0%; Score 180; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 9.1e-170; Indels 0; Gaps 0;  
Matches 180; Conservative 0; Mismatches 0;

Qy 1 MKIKLFFVTISVTISLTSITSVVLACSVTSEVHMIDDNGIKQSIGTWTFTDTPKGLQIK 60  
Db 1 MKIKLFFVTISVTISLTSITSVVLACSVTSEVHMIDDNGIKQSIGTWTFTDTPKGLQIK 60  
Qy 61 TDLKGLPAGEHGFIHEGGSCGPAEHDGHLTAGLQAHGHYDPPDKTGHEGPLGNHGKGLD 120  
Db 61 TDLKGLPAGEHGFIHEGGSCGPAEHDGHLTAGLQAHGHYDPPDKTGHEGPLGNHGKGLD 120  
Qy 121 PRLVVKADGIAKETLLAPRLTVKEIKGRTVMIHAGGDNYSKPLPLGGGGARIACGVIPN 180  
Db 121 PRLVVKADGIAKETLLAPRLTVKEIKGRTVMIHAGGDNYSKPLPLGGGGARIACGVIPN 180

RESULT 2  
ADA34763  
ID ADA34763 standard; protein; 213 AA.  
XX ADA34763;  
XX 20-NOV-2003 (first entry)  
XX Acinetobacter baumannii protein #1924.  
XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
XX plant biocontrol agent.  
XX Acinetobacter baumannii.  
XX US6562958-B1.  
XX 13-MAY-2003.  
XX 04-JUN-1999; 99US-00328352.  
XX 09-JUN-1998; 98US-0088701P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Breton G, Bush D;  
XX WPI; 2003-576092/54.  
XX N-ESDB; ADA30637.

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
PT for diagnosing a bacterial disease, as components of antibacterial  
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
PT plants.

XX Example; SEQ ID NO 6050; 328pp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.

CC The A. baumannii nucleic acids and polypeptides are useful as reagents

CC for diagnosing a bacterial disease, as components of antibacterial

CC vaccines, as targets for antibacterial drugs, to detect the presence of

CC A. baumannii and other Acinetobacter species in a sample, in screening

CC compounds for the ability to interfere with the A. baumannii life cycle

CC or to inhibit A. baumannii infection, and as biocontrol agents for

CC plants. The present sequence represents the amino acid sequence of an A.

CC baumannii protein.

XX Sequence 213 AA;

XX Query Match 8.9%; Score 16; DB 6; Length 213;

XX Best Local Similarity 100.0%; Pred. No. 3e-07; 0; Gaps 0;

XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 PLPLGGGGGARIACGVI 178

DB 197 PLPLGGGGGARIACGVI 212

RESULT 3

ABO62994

ID ABO62994 standard; protein; 144 AA.

AC ABO62994;

XX 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polypeptide segid 9511.

XX Recombinant expression vector; transcription regulatory element;

KW Klebsiella pneumoniae protein; antibacterial; vaccine.

XX Klebsiella pneumoniae.

OS US6610836-B1.

PN 26-AUG-2003.

PD 27-JAN-2000; 2000US-00489039.

PF 29-JAN-1999; 99US-0117747P.

PR (GENO-) GENOME THERAPEUTICS CORP.

PA Breton GL, Osborne M;

PI WPI; 2003-895346/82.

DR N-PSDB; ACH96545.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for

PT preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 9511; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella

CC pneumoniae polypeptide. Also described are: a recombinant expression

CC vector comprising the nucleic acid, operably linked to a transcription

CC regulatory element; and a cell comprising the recombinant expression

CC vector. The nucleic acid is useful for preparing a vaccine composition

CC against Klebsiella pneumoniae. This is the amino acid sequence of a

CC Klebsiella pneumoniae polypeptide of the invention

XX Sequence 144 AA;

XX Query Match 5.0%; Score 9; DB 7; Length 144;

XX Best Local Similarity 100.0%; Pred. No. 1.8;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 TKGHEGPLG 113

DB 69 TKGHEGPLG 77

RESULT 4

AAR32374

ID AAR32374 standard; protein; 154 AA.

XX AAR32374;

AC AAR32374;

XX 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 30-JUN-1993 (first entry)

XX 20kD Brucella abortus copper-zinc superoxide dismutase.

DE CuZnSOD; bovine brucellosis infection; detection.

XX Brucella melitensis biovar Abortus.

OS Location/Qualifiers

FH 1. .96

FT /note= "CL1"

FT 1. .50

FT /note= "N-terminal"

FT 1. .7

FT /note= "CB1"

FT 8. .37

FT /note= "CB2"

FT 38. .91

FT /note= "CB3"

FT 92. .154

FT /note= "CB4-CBS? Not clear from specification"

FT 97. .123

FT /note= "CL2"

FT 119. .138

FT /note= "SA10"

FT 124. .147

FT /note= "CL3"

FT 137. .142

FT /note= "amphipathic helix"

FT 139. .154

FT /note= "SA11"

FT 148. .154

FT /note= "CL4"

XX US5188936-A.

XX 23-FEB-1993.

XX 16-JAN-1991; 91US-00641346.

XX 16-JAN-1991; 91US-00641346.

XX (USDA ) US SEC OF AGRIC.

PA (IOWA ) UNIV IOWA STATE RES FOUND.

XX Tabatabai LB, Mayfield JE, Beck BL;

XX WPI; 1993-085536/10.

XX Detection of Brucella abortus antibody - using B. abortus copper-zinc

PT superoxidodismutase protein or segments contg. antigenic determinants.

XX Disclosure; Fig 1; 12pp; English.

XX The sequence is that of the 20kD Brucella abortus copper-zinc superoxide

CC dismutase (CuZnSOD) which is used as part of a method for detecting B.

CC abortus infection in animals, in particular Bovine brucellosis. The

CC method can distinguish between animals which have a natural infection and

CC those which have been vaccinated. CuZnSOD or a segment effective as an

CC antigenic determinant is combined with a body fluid sample and the

CC presence of a complex of the CuZnSOD and antibody detected. (Updated on  
CC 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI  
CC field.) (Updated on 24-OCT-2003 to standardise OS field)  
XX  
SQ Sequence 154 AA;  
Query Match 5.0%; Score 9; DB 2; Length 154;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 155 GGDNYSDKP 163  
Db 130 GGDNYSDKP 138  
|||||  
RESULT 5  
ID ABP82477 standard; peptide; 18 AA.  
XX  
AC ABP82477;  
DT 04-MAR-2003 (first entry)  
XX  
DE G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1150.  
XX  
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;  
KW G protein-coupled receptor modulator; antibody; immune-related disease;  
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;  
KW immunological-related cell proliferative disease; autoimmune disease;  
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;  
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;  
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;  
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;  
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;  
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;  
KW ulcer.  
XX  
OS Homo sapiens.  
XX  
XX WO200261087-A2.  
XX  
PD 08-AUG-2002.  
XX  
XX 19-DEC-2001; 2001WO-US050107.  
XX  
XX 19-DEC-2000; 2000US-0257144P.  
XX  
XX (LIFE-) LIFESPAN BIOSCIENCES INC.  
XX  
XX Burmer GC, Roush CL, Brown JP;  
XX WPI; 2003-046718/04.  
XX  
XX New isolated antigenic peptides e.g., for G protein-coupled receptors  
XX (GPCR), useful for diagnosing and designing drugs for treating conditions  
XX in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or  
XX autoimmune diseases.  
XX  
XX Claim 1; Fig 2; 523pp; English.  
XX  
XX The present invention describes antigenic peptides (I) comprising: (a)  
XX any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino  
XX acids. Also described: (1) an assay for the detection of a particular G  
XX protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;  
XX and (2) an isolated antibody having high specificity and high affinity or  
XX avidity for a particular GPCR. (I) can be used as GPCR modulators and in  
XX gene therapy. The antigenic peptides for GPCRs are useful in detecting an  
XX antibody against a particular GPCR, and in the production of specific  
XX antibodies. The peptides and antibodies are also useful for detecting the  
XX presence or absence of corresponding GPCRs. The antigenic peptides for  
XX GPCRs and antibodies are useful for diagnosing and designing drugs for  
XX treating immune-related diseases, growth-related diseases, cell  
XX regeneration-related disease, immunological-related cell proliferative

CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,  
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,  
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute  
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host  
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,  
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory  
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, ulcers, or  
CC hypotension, renal disorders, rheumatoid arthritis, trauma, hypertension,  
CC any other disorder in which GPCRs are involved. The antibodies may be  
CC used in immunoassays and immunodiagnosis. ABZ42523 to ASZ42869 encode  
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the  
CC exemplification of the present invention  
XX  
SQ Sequence 18 AA;  
Query Match 4.4%; Score 8; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 122 RLVVKADG 129  
Db 3 RLVVKADG 10  
|||||  
RESULT 6  
ABP77456  
ID ABP77456 standard; protein; 153 AA.  
XX  
AC ABP77456;  
XX  
XX 29-AUG-2003 (revised)  
DT 23-JUL-2002 (first entry)  
XX  
XX Fungi stress response protein SEQ ID NO 23.  
XX  
XX Fungi; stress response; organic acid; amino acid; fatty acid; pyrimidine;  
KW purine; nucleotide; nucleoside; lipid; diol; carbohydrate; vitamin;  
KW aromatic compound; riboflavin; food; fodder; cosmetic; pharmaceutical;  
KW genomic mapping; protein structure; bioreactor.  
XX  
XX Eremothecium gossypii.  
XX  
XX WO200222823-A2.  
XX  
XX 21-MAR-2002.  
XX  
XX 13-SEP-2001; 2001WO-EP010573.  
XX  
XX 14-SEP-2000; 2000DE-01045834.  
XX  
XX (BADI ) BASF AG.  
XX  
XX Althoefer H, Revuelta Doval JL, Santos M;  
XX WPI; 2002-383129/41.  
XX  
XX N-PSDB; ABL60138.  
XX  
XX New nucleic acid from Ashbya gossypii, useful, when modulated, for  
XX producing fine chemicals, e.g. amino acids, encodes stress-associated  
XX proteins.  
XX  
XX Claim 18; Page 108-109; 230pp; German.  
XX  
XX The invention relates to an isolated nucleic acid (I) from Ashbya  
XX gossypii comprising: (a) 63 fully defined nucleotide sequences (ABP60128-  
XX ABL60157) given in the specification; or (b) a fragment of (a). Cells  
XX that express (I) show altered production of fine chemicals, particularly:  
XX (a) organic, amino or fatty acids; (b) purine and pyrimidine bases; (c)  
XX nucleotides or nucleosides; (d) lipids; (e) diols; (f) carbohydrates; (g)  
XX aromatic compounds; (h) vitamins (particularly riboflavin); (i) co-  
XX factors; and (j) enzymes. These chemicals are useful in the food, fodder,  
XX cosmetic and pharmaceutical industries. (I) are also useful for  
XX identification and characterisation of A. gossypii and related organisms

CC and as primers and probes for detection and amplification of (I). The  
 CC encoded polypeptides (ABB77446-ABB77475) are useful as markers for  
 CC specific gene regions, in genomic mapping and for functional studies on  
 CC proteins and for evolution or protein structure investigations.  
 CC Modulating activity of (I) may improve yields, production and/or  
 CC efficiency of production of fine chemicals by *A. gossypii*. Particularly  
 CC the cells show a more robust response against stress, so viability and  
 CC productivity are improved, particularly in large-scale bioreactors.  
 CC (Updated on 29-AUG-2003 to standardise OS field)  
 XX  
 XX Sequence 153 AA;

Query Match 4.4%; Score 8; DB 5; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHGFIHE 77

Db 62 EHGFIHE 69

# RESULT 7

ADC00874  
 ID ADC00874 standard; protein; 175 AA.

XX AC ADC00874;

DT 04-DEC-2003 (first entry)

DE Enterohaemorrhagic *E. coli* O157:H7-specific protein SEQ ID NO: 919.

XX enterohaemorrhagic; anti-bacterial.

XX *Escherichia coli*; O157:H7.

PN JP2002355074-A.

XX 10-DEC-2002.

PF 24-JAN-2002; 2002JP-00015959.

PR 24-JAN-2001; 2001JP-00112010.

XX (UYTS-) UNIV TSUKUBA.

XX WPI; 2003-451640/43.

PT Enterohaemorrhagic *Escherichia coli* O157:H7-specific nucleic acid molecule  
 PT and a polypeptide and its use, a polypeptide, a vector and a host cell.

PS Claim 3; SEQ ID NO 919; 2067pp; Japanese.

CC The invention relates to a novel enterohaemorrhagic *Escherichia coli*  
 CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention  
 CC has anti-bacterial activity. The polypeptide can be used in detection  
 CC and/or treatment of O157:H7 infection. The nucleotide sequence of the  
 CC genome of Enterohaemorrhagic *E. coli* O157:H7 was determined. The present  
 CC sequence represents an *E. coli* O157:H7-specific polypeptide of the  
 CC invention.

XX Sequence 175 AA;

Query Match 4.4%; Score 8; DB 7; Length 175;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172

Db 161 PLGGGGAR 168

# RESULT 8

ADC00431

ID ADC00431 standard; protein; 175 AA.

XX AC ADC00431;

DT 04-DEC-2003 (first entry)

DE Enterohaemorrhagic *E. coli* O157:H7-specific protein SEQ ID NO: 476.

XX enterohaemorrhagic; anti-bacterial.

XX *Escherichia coli*; O157:H7.

PN JP2002355074-A.

PD 10-DEC-2002.

PF 24-JAN-2002; 2002JP-00015959.

PR 24-JAN-2001; 2001JP-00112010.

XX (UYTS-) UNIV TSUKUBA.

XX WPI; 2003-451640/43.

PT Enterohaemorrhagic *Escherichia coli* O157:H7-specific nucleic acid molecule  
 PT and a polypeptide and its use, a polypeptide, a vector and a host cell.

PS Claim 3; SEQ ID NO 476; 2067pp; Japanese.

CC The invention relates to a novel enterohaemorrhagic *Escherichia coli*  
 CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention  
 CC has anti-bacterial activity. The polypeptide can be used in detection  
 CC and/or treatment of O157:H7 infection. The nucleotide sequence of the  
 CC genome of Enterohaemorrhagic *E. coli* O157:H7 was determined. The present  
 CC sequence represents an *E. coli* O157:H7-specific polypeptide of the  
 CC invention.

XX Sequence 175 AA;

Query Match 4.4%; Score 8; DB 7; Length 175;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172

Db 161 PLGGGGAR 168

# RESULT 9

ABR57484

ID ABR57484 standard; protein; 510 AA.

XX AC ABR57484;

DT 16-SEP-2003 (first entry)

DE *Flavobacterium sfgA* protein SEQ ID NO:28.

XX *Flavobacterium*; sulphated fucogalactan digesting enzyme; polysaccharide;  
 XX sulphated fucogalactan; structural analysis; sugar engineering reagent;  
 XX sulphated fucogalactan degradation; cancer; viral infection.

OS *Flavobacterium* sp.

PN WO2003023036-A1.

XX 20-MAR-2003.

PF 05-SEP-2002; 2002WO-JP009010.

PR 05-SEP-2001; 2001JP-00268250.

XX (TAKA-) TAKARA BIO INC.

XX Ueno H, Tomono J, Sagawa H, Sakai T, Kato I;  
PI WPI; 2003-333042/31.  
DR N-PSDB; ACF03668.  
XX Sulfated fucogalactan digesting enzyme of Flavobacterium origin for  
PT polysaccharide structural analysis and engineering and preparation of  
PT degradation products.  
XX Claim 1; Page 78-82; 90pp; Japanese.  
XX The present invention describes polypeptides (P) derived from  
CC Flavobacterium sp. SA-0082 (FERM BP-5402) which have the ability to  
CC digest sulphated fucogalactans. Also described: (1) polypeptides with  
CC similar activity derived from (P) by addition, deletion and/or  
CC substitution of one or more amino acid residues, or at least 30%  
CC homologous to them; (2) nucleic acids (I) encoding the polypeptides; (3)  
CC expression vectors containing the nucleic acids; (4) hosts transformed by  
CC these vectors; (5) a method for the preparation of the polypeptides by  
CC culture of the transformed hosts; (6) sulphated fucogalactan digestion  
CC products obtained by polysaccharide digestion using the polypeptides; (7)  
CC a screening method for genes encoding polypeptides having sulphated  
CC fucogalactan digesting activity, using nucleic acids (I) or their partial  
CC sequences as probes; and (8) a method for the structural analysis of  
CC polysaccharides, using polypeptides having sulphated fucogalactan  
CC digesting activity. (P) can be used as sugar engineering reagents, for  
CC the structural analysis of polysaccharides, and for the preparation of  
CC sulphated fucogalactan degradation products, for use as antigens for the  
CC preparation of antibodies for the diagnosis of diseases including cancer  
CC and viral infection. The present sequence represents Flavobacterium sfGA,  
CC which is used in an example from the present invention  
XX  
SQ Sequence 510 AA;  
  
Query Match 4.4%; Score 8; DB 6; Length 510;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 167 GGGGARTIA 174  
Db 92 GGGGARTIA 99  
  
RESULT 10  
AAP96205  
ID AAP96205 standard; protein; 532 AA.  
XX  
AC AAP96205;  
XX  
DT 25-MAR-2003 (revised)  
DT 17-DEC-2001 (revised)  
DT 27-SEP-1989 (first entry)  
XX  
XX Human muscarinic acetylcholine m5 receptor.  
DE  
XX Muscarinic acetylcholine receptor; m5; drug screening; diagnosis;  
KW drug design.  
KW  
XX Homo sapiens.  
OS  
XX USN7241971-N.  
FN  
XX 14-MAR-1989.  
PD  
XX 08-SEP-1988; 88US-00241971.  
PF  
XX 08-SEP-1988; 88US-00241971.  
PR  
XX (USSH) NAT INST OF HEALTH.  
PA  
XX Bonner T;  
PI  
XX

DR WPI; 1989-165452/22.  
DR N-PSDB; AAN92071.  
XX Cloned genes for muscarinic acetylcholine receptors - for drug screening  
PT and diagnostic use.  
PT  
XX Disclosure; Page; 21pp; English.  
XX  
XX Cell lines producing this MAR is useful in screening for drugs with  
CC improved specificity for the part. MAR type. The amino acid sequences may  
CC be useful for rational drug design and antibodies to the sequences may be  
CC useful for diagnostic purposes. See also AAN92068-N92067 and AAN90086.  
CC (Note: Revised entry submitted to correct the patent number format of US  
CC Government-owned NTIS applications to prevent clashes with ongoing US  
CC granted patent numbers. For further information please visit the Derwent  
CC web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).) (Updated on 25-  
CC MAR-2003 to correct PI field.)  
XX  
SQ Sequence 532 AA;  
  
Query Match 4.4%; Score 8; DB 1; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 122 RLNVKADG 129  
Db 382 RLNVKADG 389  
|||||||  
  
RESULT 11  
ABB56364  
ID ABB56364 standard; protein; 532 AA.  
XX  
AC ABB56364;  
XX  
XX 18-FEB-2002 (first entry)  
DT  
XX Non-endogenous human GPCR protein, SEQ ID NO: 521.  
DE  
XX Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;  
KW constitutively activated GPCR; agonist; disease.  
KW  
XX Homo sapiens.  
OS  
XX Synthetic.  
OS  
XX WO200177172-A2.  
FN  
XX 18-OCT-2001.  
PD  
XX  
XX 05-APR-2001; 2001WO-US011098.  
PF  
XX  
XX 07-APR-2000; 2000US-0195747P.  
PR  
XX  
XX (AREN-) ARENA PHARM INC.  
PA  
XX  
XX Lehmann-Bruinsma K, Liaw CW, Lin I;  
PI  
XX WPI; 2001-648759/74.  
KW  
XX N-PSDB; ABI98000.  
DR  
XX  
PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in  
PT disease treatment, comprises contacting candidate compounds with versions  
PT of GPCRs.  
XX  
XX Claim 1; Page 322-324; 394pp; English.  
PS  
XX The invention relates to G protein-coupled receptors (GPCRs) for which  
CC the endogenous ligand has been identified. Non-endogenous constitutively  
CC activated versions of known GPCRs are used in the invention for the  
CC direct identification of candidate compounds as receptor agonists,  
CC inverse agonists or partial agonists. Such agonists are useful as  
CC therapeutic agents for diseases or disorders associated with GPCRs. The  
CC present sequence is a non-endogenous version of a known human GPCR

XX SQ Sequence 532 AA;  
Query Match 4.4%; Score 8; DB 4; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLVVKADG 129  
|||  
Db 382 RLVVKADG 389

RESULT 12  
AAU97746  
ID AAY97746 standard; protein; 532 AA.  
AC AAY97746;  
XX  
DT 12-JUL-2001 (first entry)  
XX  
DE Human CHRM5 protein sequence #1.  
XX  
KW CHRM5; human; cholinergic receptor muscarinic 5; polymorphic variant;  
KW genotyping; haplotype; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200128995-A2.  
XX  
PD 26-APR-2001.  
XX  
PF 19-OCT-2000; 2000WO-US029071.  
XX  
PR 21-OCT-1999; 99US-0160647P.  
XX  
PA (GENA-) GENAISSANCE PHARM INC.  
XX  
PI Chew A, Choi JY, Nandabalan K, Stephens JC;  
XX WPI: 2001-300313/31.  
XX N-PSDB; AA921430.

Isolated polynucleotide encoding the cholinergic receptor, muscarinic 5 (CHRM5), used to genotype/haplotype the CHRM5 gene, and to identify an association between a trait and a polymorphism, comprises novel polymorphisms.

PS Disclosure; Fig 3; 53pp; English.

CC This sequence is the human cholinergic receptor, muscarinic 5 (CHRM5) protein. The invention relates to a polymorphic variant of the CHRM5 gene sequence. The polymorphic sequence is useful to genotype or haplotype the CHRM5 gene, to predict a haplotype pair for the CHRM5 gene, and for identifying an association between a trait (such as a clinical response to a drug targeting CHRM5). It is also useful in gene therapy in patients who lack the CHRM5 isogene or have only one copy of it, and in assays to measure the binding affinities of one or more candidate drugs targeting CHRM5. The DNA sequence is used in the treatment of disorders affected by expression or function of a novel CHRM5 isogene of the invention. The protein encoded by the CHRM5 variant is useful to identify drugs which target the CHRM5 polymorphic variant protein. Antibodies against the protein can be used to neutralise the CHRM5 isoform activity expressed in an individual, and is useful in detection of CHRM5 in immunocytochemical, immunohistochemical and immunofluorescence. A composition containing a genotyping oligonucleotide for detecting a polymorphism in the CHRM5 gene is used to detect novel CHRM5 polymorphisms of the invention

XX SQ Sequence 532 AA;  
Query Match 4.4%; Score 8; DB 4; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLVVKADG 129  
|||  
Db 382 RLVVKADG 389

RESULT 13  
AAU97549  
ID AAU97549 standard; protein; 532 AA.  
XX  
AC AAU97549;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Human cholinergic receptor, muscarinic 5 (CHRM5) protein.  
XX  
KW Human; cholinergic receptor muscarinic 5; CHRM5; genotyping; haplotyping;  
KW single nucleotide polymorphism; SNP; chromosome 15q26.  
XX  
OS Homo sapiens.  
XX  
PN WO200232924-A2.  
XX  
PD 25-APR-2002.  
XX  
PF 11-OCT-2001; 2001WO-US032022.  
XX  
PR 19-OCT-2000; 2000WO-US029071.  
XX (GENA-) GENAISSANCE PHARM INC.  
XX  
PI Bieglecki KM, Chew A, Choi JY, Denton RR, Nandabalan K;  
PI Sausker EA, Stephens JC;  
XX  
XX WPI: 2002-435523/46.  
XX N-PSDB; ABX52224, ABX52225.

Novel cholinergic receptor, muscarinic 5 polynucleotide useful therapeutically and in screening for candidate drug to treat diseases related to the receptor activity.

XX Claim 27; Fig 3; 72pp; English.

CC The present invention relates to a new cholinergic receptor, muscarinic 5 (CHRM5) polynucleotide comprising a sequence which is a polymorphic variant for a reference sequence for the CHRM5 gene or its fragment, or a polymorphic variant of a reference sequence for a CHRM5 cDNA or its fragment. The invention is useful in drug screening assays. The molecules of the invention are useful in studying the expression and function of CHRM5, and in expressing CHRM5 protein for use in screening for candidate drugs to treat diseases related to CHRM5 activity. The methods of the invention are useful in developing diagnostic tests and therapeutic treatments. The method is also useful in the design of clinical trials of candidate drugs for treating specific condition or disease associated with CHRM5 activity and is useful in determining whether an individual has one of the haplotypes or one of the haplotype pairs. The invention is useful in a variety of diagnostic and prognostic formats and therapeutic methods. The invention is also useful in genotyping and/or haplotyping the CHRM5 gene in an individual. The present amino acid sequence represents the human CHRM5 protein of the invention. This sequence is encoded by the human CHRM5 gene located on chromosome 15q26

XX SQ Sequence 532 AA;  
Query Match 4.4%; Score 8; DB 5; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLVVKADG 129  
|||  
Db 382 RLVVKADG 389

RESULT 14

| Qy  | Db       | Query Match | Best Local Similarity | Matches | Conservative          | Mismatches | Indels | Gaps | Length |
|-----|----------|-------------|-----------------------|---------|-----------------------|------------|--------|------|--------|
| 122 | RLVVKADG | 129         | 4.4%; Score 8;        | DB 5;   | 100.0%; Pred. No. 60; | 0;         | 0;     | 0;   | 532;   |
| 382 | RLVVKADG | 389         | 8;                    | 0;      | 0;                    | 0;         | 0;     | 0;   | 532;   |



Query Match 4.4%; Score 8; DB 6; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLVRKADG 129  
DB 382 RLVRKADG 389

## RESULT 16

ADD29411  
ID ADD29411 standard; protein; 532 AA.

XX AC ADD29411;  
XX 15-JAN-2004 (first entry)  
XX Human muscarinic acetylcholine receptor M5 protein.

DE XX metabolic disorder; muscarinic acetylcholine receptor; M5 modulator;  
KW acetylcholine signalling pathway; metabolic function; immunomodulator;  
KW anorectic; anabolic; aberrant food intake; obesity; cachexia; anorexia;  
KW human; M5.

XX Homo sapiens.  
XX US2003092041-A1.  
XX 15-MAY-2003.

XX 22-AUG-2002; 2002US-00225928.  
XX 23-AUG-2001; 2001US-0314324P.  
XX (MILL-) MILLENNIUM PHARM INC.

XX White DW;

XX WPI; 2003-765476/72.  
XX N-PSDB; ADD29410, ADD29412.

PT Treating a subject having a metabolic disorder exhibiting muscarinic  
PT receptor polypeptide activity or aberrant M5 nucleic acid expression,  
PT comprises administering M5 modulator.

XX Claim 27; SEQ ID NO 2; 34pp; English.

XX This invention relates to a novel method for the treatment of a subject  
CC having a metabolic disorder exhibiting aberrant muscarinic acetylcholine  
CC receptor M5 polypeptide activity or aberrant M5 nucleic acid expression  
CC which comprises administering to the subject an M5 modulator. M5  
CC nucleotides and polypeptides play a role in or function in acetylcholine  
CC signalling pathways which are involved in regulation of metabolic  
CC function. The M5 modulators of the invention may have immunomodulator,  
CC anorectic or anabolic activity. The invention is useful for identifying a  
CC compound capable of treating a metabolic disorder such as aberrant food  
CC intake, obesity, cachexia, or anorexia. The present sequence is that of  
CC the human M5 protein which was used during the identification of the M5  
CC modulators of the invention.

XX Sequence 532 AA;

Query Match 4.4%; Score 8; DB 7; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLVRKADG 129  
DB 382 RLVRKADG 389

## RESULT 17

ADE40455

ID ADE40455 standard; protein; 532 AA.

XX AC ADE40455;

XX 29-JAN-2004 (first entry)

XX Human muscarinic acetylcholine receptor M5 (gene ID 126) protein.

XX AIDS; acquired immunodeficiency syndrome; human immunodeficiency virus;  
KW HIV-related disorder; differential expression; drug screening;  
KW viral replication modulation; diagnosis; prognosis; predisposition;  
KW anti-HIV; gene therapy; antisense therapy; human;  
KW muscarinic acetylcholine receptor M5; receptor.

XX Homo sapiens.

XX WO2003070883-A2.

XX 28-AUG-2003.

XX 13-FEB-2003; 2003WO-US004246.

XX 15-FEB-2002; 2002US-0357391P.

XX 13-MAY-2002; 2002US-0380249P.

XX 25-JUN-2002; 2002US-0391306P.

XX 27-AUG-2002; 2002US-0406297P.

XX 19-SEP-2002; 2002US-0412007P.

XX 10-OCT-2002; 2002US-0417508P.

XX 10-DEC-2002; 2002US-0432318P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Powell DM, Weich NS;

XX WPI; 2003-671808/63.

XX N-PSDB; ADE40454.

XX Identifying a compound capable of diagnosing, preventing or treating AIDS  
PT or an HIV-related disorder comprises assaying the ability of the compound  
PT to modulate e.g. 1414, 1481 or 1553 nucleic acid expression or  
PT polypeptide activity.

XX Claim 1; SEQ ID NO 34; 167pp; English.

XX The invention relates to a method of identifying a compound useful in the  
CC treatment of AIDS (acquired immunodeficiency syndrome) or an HIV (human  
CC immunodeficiency virus)-related disorder. The invention involves assaying  
CC the ability of a test compound to modulate the activity or expression of  
CC 26 human proteins. These proteins and nucleic acids encoding them  
CC (ADE40422-ADE40473) are differentially expressed in tissues relating to  
CC AIDS or an HIV-related disorder compared to their expression in normal  
CC tissues. The invention also relates to the use of the compounds  
CC identified to modulate viral replication in a cell and to treat a patient  
CC with AIDS or an HIV-related disorder. The invention further discloses  
CC methods for the diagnostic evaluation and prognosis of various HIV-  
CC related disorders, and for the identification of individuals exhibiting a  
CC predisposition to such conditions. The modulatory compounds identified  
CC using the method of the invention may be small organic molecules,  
CC peptides, antibodies or antisense nucleic acid molecules. The methods of  
CC the invention are useful in diagnosing, preventing or treating AIDS or  
CC HIV-related disorders. The present sequence represents a human protein  
CC which is differentially expressed in AIDS or HIV-related disorders.

XX Sequence 532 AA;

Query Match 4.4%; Score 8; DB 7; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLVRKADG 129  
DB 382 RLVRKADG 389

RESULT 18  
ADO29257  
ID ADO29257 standard; protein; 532 AA.  
XX  
AC ADO29257;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human GPCR CHRMS, SEQ ID NO:358.  
XX  
XX G protein-coupled receptor; GPCR; drug screening; diagnosis;  
KW transgenic mouse; neurological disorder; adrenal gland disorder;  
KW colon disorder; intestinal disorder; cardiovascular disorder;  
KW muscular disorder; blood disorder; immune disorder; bone disorder;  
KW joint disorder; metabolic disorder; nutritive disorder; cancer;  
KW kidney disorder; liver disorder; lung disorder; breast disorder;  
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;  
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;  
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;  
KW cytosatic; antinflammatory; vasotropic; antianginal; antiarrhythmic;  
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;  
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;  
KW dermatological; antiulcer; antithyroid; antiallergic; anorectic;  
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;  
receptor.  
XX  
XX Homo sapiens.  
OS  
XX  
FN WO2004040000-A2.  
XX  
PD 13-MAY-2004.  
XX  
PF 09-SEP-2003; 2003WO-US028226.  
XX  
PR 09-SEP-2002; 2002US-0409303P.  
PR 09-APR-2003; 2003US-0461329P.  
XX  
PA (PRIM-) PRIMAL INC.  
XX  
PI Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;  
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;  
XX  
DR WPI; 2004-390329/36.  
DR N-PSDB; ADO29847.  
XX  
XX Novel mammalian G protein coupled receptors, useful for identifying  
PT compounds that modulates diagnosing and treating disease condition  
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina  
PT pectoris, Parkinson's disease.  
XX  
XX Claim 151; SEQ ID NO 358; 542pp; English.  
XX  
XX The invention relates to human and mouse G protein-coupled receptors  
CC (GPCRs) and nucleic acids encoding them. The invention also relates to  
CC sequences at least 90% identical to the GPCR proteins and nucleic acids  
CC of the invention; methods of treating, preventing or diagnosing diseases  
CC associated with GPCRs of the invention; methods of screening for  
CC compounds useful in the treatment of GPCR-related diseases; a transgenic  
CC mouse comprising a GPCR gene of the invention; a mouse comprising a  
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived  
CC from the transgenic mice; kits comprising several mice, each of which has  
CC a mutation in a different GPCR gene of the invention; and kits comprising  
CC probes which hybridise to GPCR polynucleotides of the invention. The  
CC invention further discloses variants of the GPCR polypeptides and vectors  
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may  
CC be used in the diagnosis, treatment or prevention of a wide variety of  
CC diseases including neurological disorders (e.g., Alzheimer's disease,  
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);  
CC disorders of the adrenal gland; disorders of the colon or intestine  
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel  
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or

myocardial infarction); muscular disorders; blood disorders (e.g.,  
anaemia or leukaemia); immune disorders (e.g., osteoarthritis, rheumatoid  
ADDS); bone and joint disorders (e.g., osteoporosis); metabolic or nutritive disorders (e.g.,  
arthritis, gout or osteoporosis); enzyme deficiency-related diseases or vitamin deficiency-related  
diseases); and disorders of the kidney, liver, lung, breast, ovary,  
uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and  
thyroid (e.g., cancers). The present sequence represents a GPCR of the  
invention. Note: The full sequence data for this patent did not form part  
of the printed specification; those sequences not shown were obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 532 AA;  
Query Match 4.4%; Score 8; DB 8; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 122 RLVVVKADG 129  
Db 382 RLVVVKADG 389  
|||||||  
RESULT 19  
ADO29258  
ID ADO29258 standard; protein; 532 AA.  
XX  
AC ADO29258;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Mouse GPCR CHRMS, SEQ ID NO:359.  
XX  
KW G protein-coupled receptor; GPCR; drug screening; diagnosis;  
KW transgenic mouse; neurological disorder; adrenal gland disorder;  
KW colon disorder; intestinal disorder; cardiovascular disorder;  
KW muscular disorder; blood disorder; immune disorder; bone disorder;  
KW joint disorder; metabolic disorder; nutritive disorder; cancer;  
KW kidney disorder; liver disorder; lung disorder; breast disorder;  
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;  
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;  
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;  
KW cytosatic; antinflammatory; vasotropic; antianginal; antiarrhythmic;  
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;  
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;  
KW dermatological; antiulcer; antithyroid; antiallergic; anorectic;  
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;  
murine; receptor.  
XX  
XX Mus musculus.  
OS  
XX  
FN WO2004040000-A2.  
XX  
PD 13-MAY-2004.  
XX  
PF 09-SEP-2003; 2003WO-US028226.  
XX  
PR 09-SEP-2002; 2002US-0409303P.  
PR 09-APR-2003; 2003US-0461329P.  
XX  
PA (PRIM-) PRIMAL INC.  
XX  
PI Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;  
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;  
XX  
DR WPI; 2004-390329/36.  
DR N-PSDB; ADO30138.  
XX  
XX Novel mammalian G protein coupled receptors, useful for identifying  
PT compounds that modulates diagnosing and treating disease condition  
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina  
PT pectoris, Parkinson's disease.

XX Claim 151; SEQ ID NO 359; 542pp; English.  
XX The invention relates to human and mouse G protein-coupled receptors  
XX (GPCRs) and nucleic acids encoding them. The invention also relates to  
XX sequences at least 90% identical to the GPCR proteins and nucleic acids  
XX of the invention; methods of treating, preventing or diagnosing diseases  
XX associated with GPCRs of the invention; methods of screening for  
XX compounds useful in the treatment of GPCR-related diseases; a transgenic  
XX mouse comprising a GPCR gene of the invention; a mouse comprising a  
XX mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived  
XX from the transgenic mice; kits comprising several mice, each of which has  
XX a mutation in a different GPCR gene of the invention; and kits comprising  
XX probes which hybridise to GPCR polynucleotides of the invention. The  
XX invention further discloses variants of the GPCR polypeptides and vectors  
XX comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may  
XX be used in the diagnosis, treatment or prevention of a wide variety of  
XX diseases including neurological disorders (e.g., Alzheimer's disease,  
XX depression, diabetic neuropathy, Parkinson's disease or schizophrenia);  
XX disorders of the adrenal gland; disorders of the colon or intestine  
XX (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel  
XX syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or  
XX myocardial infarction); muscular disorders; blood disorders (e.g.,  
XX anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or  
XX AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid  
XX arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,  
XX obesity, enzyme deficiency-related diseases or vitamin deficiency-related  
XX diseases); and disorders of the kidney, liver, lung, breast, ovary,  
XX uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and  
XX thyroid (e.g., cancers). The present sequence represents a GPCR of the  
XX invention. Note: The full sequence data for this patent did not form part  
XX of the printed specification; those sequences not shown were obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 532 AA;

Query Match 4.4%; Score 8; DB 8; Length 532;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 122 RLVRKADG 129  
DB 382 RLVRKADG 389  
|||||

RESULT 20  
ABR57482  
ID ABR57482 standard; protein; 533 AA.  
XX ABR57482;  
XX 16-SEP-2003 (first entry)  
XX Flavobacterium sfgA protein SEQ ID NO:20.  
XX Flavobacterium; sulphated fucogalactan digesting enzyme; polysaccharide;  
XX sulphated fucogalactan; structural analysis; sugar engineering reagent;  
XX sulphated fucogalactan degradation; cancer; viral infection.  
XX Flavobacterium sp.  
XX WO2003023036-A1.  
XX 20-MAR-2003.  
XX 05-SEP-2002; 2002WO-JF009010.  
XX 05-SEP-2001; 2001JP-00268250.  
XX (TAKA-) TAKARA BIO INC.  
XX Ueno H, Tomono J, Sagawa H, Sakai T, Kato I;

XX WPI; 2003-333042/31.  
XX N-PSDB; ACF03662.  
XX Sulphated fucogalactan digesting enzyme of Flavobacterium origin for  
XX polysaccharide structural analysis and engineering and preparation of  
XX degradation products.  
XX Example 1; Page 67-70; 90pp; Japanese.  
XX The present invention describes polypeptides (P) derived from  
XX Flavobacterium sp. SA-0082 (FERM BP-5402) which have the ability to  
XX digest sulphated fucogalactans. Also described: (1) polypeptides with  
XX similar activity derived from (P) by addition, deletion and/or  
XX substitution of one or more amino acid residues, or at least 30%  
XX homologous to them; (2) nucleic acids (I) encoding the polypeptides; (3)  
XX expression vectors containing the nucleic acids; (4) hosts transformed by  
XX these vectors; (5) a method for the preparation of the polypeptides by  
XX culture of the transformed hosts; (6) sulphated fucogalactan digestion  
XX products obtained by polysaccharide digestion using the polypeptides; (7)  
XX a screening method for genes encoding polypeptides having sulphated  
XX fucogalactan digesting activity, using nucleic acids (I) or their partial  
XX sequences as probes; and (8) a method for the structural analysis of  
XX polysaccharides, using polypeptides having sulphated fucogalactan  
XX digesting activity. (P) can be used as sugar engineering reagents, for  
XX the structural analysis of polysaccharides, and for the preparation of  
XX sulphated fucogalactan degradation products for use as antigens for the  
XX preparation of antibodies for the diagnosis of diseases including cancer  
XX and viral infection. The present sequence represents Flavobacterium sfgA,  
XX which is used in an example from the present invention  
XX  
SQ Sequence 533 AA;  
Query Match 4.4%; Score 8; DB 6; Length 533;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 167 GGGGARIA 174  
DB 115 GGGGARIA 122  
|||||  
RESULT 21  
ADC37562  
ID ADC37562 standard; protein; 684 AA.  
XX ADC37562;  
XX 18-DEC-2003 (first entry)  
XX Human nucleic acid associated protein, NAAP-29.  
XX Human; nucleic acid associated protein; NAAP; cytostatic;  
XX antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;  
XX cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;  
XX thymometric; gene therapy; cell proliferative disorder; cancer;  
XX atherosclerosis; neurological disorder; epilepsy; Huntington's disease;  
XX stroke; immune disorder; inflammatory disorder; AIDS; allergy;  
XX developmental disorder; Hypothyroidism; Cushing's syndrome; infection.  
XX Homo sapiens.  
XX WO2003046151-A2.  
XX 05-JUN-2003.  
XX 26-NOV-2002; 2002WO-US038445.  
XX 27-NOV-2001; 2001US-0333925P.  
XX 07-DEC-2001; 2001US-0340477P.  
XX 14-DEC-2001; 2001US-0340362P.  
XX 18-DEC-2001; 2001US-0342002P.  
XX

```
(INCYTE GENOMICS INC.
XX
XX Baughn MR, Becha SD, Bhatia U, Blake JJ, Burford N, Burrill JD;
PI Chang H, Chawla NK, Elliott VS, Emerling BW, Forsythe TJ, Gandhi AR;
PI Gietzen KJ, Gorvad AE, Griffin JA, Hafalia AJA, Jackson JL, Ho A;
PI Ison CH, Jackson AA, Jiang X, Jin P, Kable AE, Khare R, Lal PG;
PI Lee EA, Lee S, Lee SY, Li JX, Lu DAM, Ramkumar J, Richardson JW;
PI Sprague WM, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
PI Zheng W;
XX
XX WPI; 2003-513642/48.
DR N-PSDB; ADC37622.
XX
XX New human nucleic acid associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
PT infections.
XX
XX Claim 1; SEQ ID NO 29; 383pp; English.
XX
XX The present invention relates to novel human nucleic acid associated
CC protein (NAAP) (I; ADC37534-ADC37593) and their coding sequences {}. The
CC NAAPs and their coding sequences are useful in diagnosing, treating and
CC preventing diseases or conditions associated with the decreased
CC expression or over expression of NAAP, such as cell proliferative (e.g.
CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
CC disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and
CC developmental (e.g. Hypothyroidism, Cushing's syndrome) disorders, or
CC infections. These are also useful in assessing the effects of exogenous
CC compounds on the expression of nucleic acid and amino acid sequences of
CC NAAP.
XX
XX Sequence 684 AA;
SQ
Query Match 4.4%; Score 8; DB 7; Length 684;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 143 KEIKGRTV 150
Db 177 KEIKGRTV 184
|||||
177 KEIKGRTV 184
RESULT 22
AAW41927
ID AAW41927 standard; protein; 759 AA.
XX
XX AAW41927;
AC
XX
XX 21-JUL-1998 (first entry)
DT
XX
XX Homo sapiens telomerase protein p105.
DE
XX
XX telomerase; p105; treatment; prevention; cancer; restenosis;
KW inflammation; myocardial infarction; glomerulonephritis; transplant;
KW rejection; infection; HIV; human immunodeficiency virus;
KW bone marrow transplants; proliferation-restricted cells.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Domain 5..81
FT Domain /note= "telomerase binding domain"
FT Domain 115..192
FT Domain /note= "telomerase binding domain"
FT Domain 336..420
FT Domain /note= "telomerase binding domain"
FT Domain 487..578
FT Domain /note= "telomerase binding domain"
XX
XX WO9801543-A1.
PN
XX
XX 15-JAN-1998.
PD
XX
XX 08-JUL-1997; 97WO-US012297.
XX
XX 08-JUL-1996; 96US-00676967.
XX (TULA-) TULARIK INC.
PA
XX
XX Cao Z;
PI
XX
XX WPI; 1998-101044/09.
DR N-PSDB; AAV13832.
XX
XX New nucleic acid encoding human telomerase protein p105 or its fragments
PT - used for therapeutic modulation of telomerase activity and for
PT screening for potential modulators of telomerase-target binding.
XX
XX Claim 1; Page 17-19; 32pp; English.
XX
XX The sequence is that of telomerase protein p105, it can be used to screen
CC for agents, e.g. antibodies, that modulate binding of human telomerase to
CC its binding target. Those that inhibit telomerase activity can be used to
CC treat conditions such as cancer, restenosis, inflammation, myocardial
CC infarction, glomerulonephritis, transplant rejection and infections (e.g.
CC with human immunodeficiency virus), while those that are agonists can be
CC used to extend the life of proliferation-restricted cells, especially
CC normal somatic cells, e.g. in cases of hypersensitivity or atrophy, also
CC to improve production of recombinant proteins by maximising cell density
CC and survival and expansion of precursor cells being used for bone marrow
CC transplants. They may also be used for diagnosis. Other uses of
CC telomerase proteins are isolation, enrichment and concentration of
CC telomerase RNA or proteins; as immunogens; in therapy; as reagent where
CC nascent oligonucleotides of known structure are needed (e.g. for tagging
CC native nucleic acid molecules) and for regulating cell growth/density
CC tolerance. The agents and the telomerase proteins should be very
CC specific, e.g. they are selective for cancer cells without harming
CC somatic cells
XX
XX Sequence 759 AA;
SQ
Query Match 4.4%; Score 8; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 143 KEIKGRTV 150
Db 177 KEIKGRTV 184
|||||
177 KEIKGRTV 184
RESULT 23
AAW46593
ID AAW46593 standard; protein; 759 AA.
XX
XX AAW46593;
AC
XX
XX 06-JUL-1998 (first entry)
DT
XX
XX Human telomerase p105 subunit.
DE
XX
XX Telomerase; p105; human; cell replication; cancer; restenosis;
KW multiple sclerosis; inflammation; glomerulonephritis; rheumatoid arthritis;
KW myocardial infarction; glomerulonephritis; transplant rejection;
KW infection; therapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Domain 5..81
FT Domain /note= "RRM1 binding domain"
FT Domain 115..192
FT Domain /note= "RRM2 binding domain"
FT Domain 336..420
FT Domain /note= "RRM3 binding domain"
FT Domain 487..578
FT Domain
```

FT XX /note= "RRMA binding domain"

PN WO9801542-A1.

XX 15-JAN-1998.

XX 08-JUL-1997; 97WO-US012296.

XX 08-JUL-1996; 96US-00676974.

XX (REGC ) UNIV CALIFORNIA.

XX Collins K;

XX WPI; 1998-101043/09.

DR N-PSDB; AAV05369, AAV05370, AAV05371, AAV05372.

XX New nucleic acid encoding human telomerase proteins or their fragments -

PT useful for therapeutic modulation of telomerase activity and for

PT screening for potential modulators of telomerase-target binding.

XX Claim 4; Page 17-19; 32pp; English.

PS This protein comprises the p105 subunit of human telomerase. p105 can be

XX isolated from human cells or expressed in host cells using native p105

CC cDNA (see AAV05369), or optimised synthetic sequences (see AAV05370-72).

CC The invention provides methods relating to human telomerase and related

CC nucleic acids, including the subunit proteins p140, p105, p48 and p43.

CC The invention also provides isolated telomerase hybridisation probes and

CC primers capable of specifically hybridising with the telomerase gene,

CC telomerase- specific binding agents such as specific antibodies, and

CC methods of making and using the subject compositions in diagnosis (e.g.

CC genetic hybridisation screens for telomerase transcripts), therapy (e.g.

CC gene therapy to modulate telomerase gene expression) and in the

CC biopharmaceutical industry (e.g. reagents for screening chemical

CC libraries for lead agents). Modulation of telomerase expression can be

CC used for the treatment or prevention of cancer, restenosis, inflammation,

CC myocardial infarction, glomerulonephritis, transplant rejection or

CC infections (e.g. with HIV). Telomerase proteins can also be used in the

CC isolation, enrichment and concentration of telomerase RNA proteins, as

CC immunogens, in therapy, for regulating cell growth/density tolerance and

CC for polymerising nucleic acid on a substrate

XX Sequence 759 AA;

SQ Query Match 4.4%; Score 8; DB 2; Length 759;

Best Local Similarity 100.0%; Pred. No. 84;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 KEIKGRTV 150

DB 177 KEIKGRTV 184

|||||

RESULT 24

AA092754

ID AAB92754 standard; protein; 759 AA.

XX AAB92754;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:11212.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PT length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length

PT cDNAs.

XX Claim 8; SEQ ID NO 11212; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-

CC length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the

CC complementary strand of a polynucleotide which comprises one of the 5602

CC nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893

CC represent human amino acid sequences; and AAH13629 to AAH13632 represent

CC oligonucleotides, all of which are used in the exemplification of the

CC present invention

XX Sequence 759 AA;

SQ Query Match 4.4%; Score 8; DB 4; Length 759;

Best Local Similarity 100.0%; Pred. No. 84;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 KEIKGRTV 150

DB 177 KEIKGRTV 184

|||||

RESULT 25

AB053053

ID AB053053 standard; protein; 956 AA.

XX AB053053;

XX 10-OCT-2003 (first entry)

XX Human putative spliceosome associated protein (SAP) #29.

XX Human; SAP; spliceosome associated protein; ribonucleoprotein;

XX RNP complex; RNA affinity substrate; RNP assembly sequence;

XX spliceosomal complex; hnRNP complex; mRNA export complex;

XX mRNA localisation complex; RNA editing complex; intron complex;

XX H complex; telomerase complex; fragile X protein complex;

XX reverse transcriptase complex; gene splicing complex.

XX Homo sapiens.

OS

```

XX US2003069803-A1.
PN
XX
XX 10-APR-2003.
PD
XX
XX 14-JAN-2002; 2002US-00047991.
PF
XX
XX 12-JAN-2001; 2001US-0261521P.
PR
XX (REED// REED R.
PA (ZHOU// ZHOU Z.
XX
XX Reed R, Zhou Z;
PI
XX
XX WPI; 2003-540885/51.
DR
XX
XX Isolating ribonucleoprotein complex, by contacting RNA affinity substrate
PT having ribonucleoprotein assembly sequence and affinity tag, with protein
PT mixture, subjecting complex formed to chromatography, affinity selection.
XX
XX Claim 24; Page; 39pp; English.
PS
XX
XX The invention relates to forming (M1) an isolated ribonucleoprotein (RNP)
CC complex (C), involves contacting an RNA affinity substrate (S) comprising
CC an RNP assembly sequence (AS) and an affinity tag, with a protein mixture
CC to permit formation of (C) on AS, subjecting (C) to chromatographic
CC separation, and subjecting (C) to affinity selection, where the affinity
CC tag (e.g. bacteriophage M2 coat protein in a fusion protein with E. coli
CC maltose binding protein) binds to an affinity matrix. Also included are
CC an isolated spliceosome preparation (isolated by (M1)), a RNA comprising
CC an RNP complex binding site and at least one phage coat protein
CC recognition site, a nucleic acid encoding the RNA, and treating (M2) a
CC subject having a disorder associated with abnormal RNP complexes (by
CC obtaining a sample of cells from a subject, purifying RNP complexes from
CC the cells of the subject by (M1), determining the presence in the
CC purified RNP complexes of one or more proteins, and normalising the
CC amount of RNPs in the subject. (M1) is useful for forming an isolated RNP
CC complex selected from a spliceosomal complex (selected from E, A, B and C
CC complex), an hnRNP complex, an mRNA export complex, an mRNA localisation
CC complex, an RNA editing complex, an intron complex, or an H complex. (M1)
CC is useful in a diagnostic assay for determining whether a subject has
CC abnormal RNP complexes, (M2) is useful for treating a subject having a
CC disorder associated with abnormal RNP complexes. (M1) is useful for
CC forming an isolated RNP complex such as a telomerase complex, a fragile X
CC protein complex, a reverse transcriptase complex or a gene splicing
CC complex. The present sequence represents a putative novel human
CC spliceosome associated protein (SAP) isolated by the methods of the
CC invention. Note: The present sequence is not shown in the specification
CC but was obtained from Genbank or Swissprot using the information provided
CC in table 2 of the specification
XX
XX Sequence 956 AA;
SQ
Query Match 4.4%; Score 8; DB 6; Length 956;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 163 PLPLGGGG 170
Db 32 PLPLGGGG 39
RESULT 26
AAG86501
ID AAG86501 standard; peptide; 10 AA.
XX
XX AAG86501;
AC
XX
XX 11-SEP-2001 (first entry)
DT
XX
XX Saccharomyces cerevisiae peptide, SEQ ID NO: 1450.
DE
XX Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW

```

---

```

KW drug discovery; drug design.
XX
XX Saccharomyces cerevisiae.
XX
XX WO200142276-A1.
PN
XX
XX 14-JUN-2001.
PD
XX
XX 13-DEC-2000; 2000WO-GB004773.
PF
XX
XX 13-DEC-1999; 99GB-00029471.
PR
XX (PROT-) PROTEOM LTD.
PA
XX
XX Roberts GW, Heal JR;
PI
XX
XX WPI; 2001-367863/38.
DR
XX
XX Identifying complementary peptides by analysis of protein and nucleotide
PT sequence databases, useful in drug design.
PT
XX
XX Example 3; Page 230; 488pp; English.
PS
XX
XX The invention relates to the identification of complementary peptides by
CC analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents and
CC drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC Saccharomyces cerevisiae
XX
XX Sequence 10 AA;
SQ
Query Match 3.9%; Score 7; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 166 LGGGGAR 172
Db 1 LGGGGAR 7
RESULT 27
ABB73398
ID ABB73398 standard; peptide; 20 AA.
XX
XX ABB73398;
AC
XX
XX 05-APR-2002 (first entry)
DT
XX
XX IL-1 R antagonist peptide SEQ ID NO:1048.
DE
XX
XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TWP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antinaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
XX Homo sapiens.
XX
XX Synthetic.
OS
XX
XX WO200183525-A2.
PN
XX
XX 08-NOV-2001.
PD
XX
XX

```

PF 02-MAY-2001; 2001WO-US014310.  
XX  
PR 03-MAY-2000; 2000US-00563286.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JW;  
XX  
XX WPI; 2002-130313/17.  
DR  
XX  
XX Novel vehicle-peptide molecule or its multimers useful for treating  
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
PT diabetic retinopathy, obesity, sleep disorders and infertility.  
XX  
XX Disclosure; Page 90; 176pp; English.  
PS  
XX  
XX The present invention describes a vehicle-peptide molecule (I) or its  
CC multimers. (I) can have antiinflammatory, antitumor, immunosuppressive,  
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and  
CC neuroprotective activities. (I) can be used as a therapeutic or  
CC prophylactic agent as well as for screening purposes. (I) is useful for  
CC diagnosing diseases characterised by dysfunction of their associated  
CC protein of interest, for identifying normal or abnormal proteins of  
CC interest, as a part of diagnostic kit to detect the presence of their  
CC proteins of interest in a biological sample. Additionally, (I) is useful  
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
CC infertility, and neurological degenerative diseases. (I), comprising EPO-  
CC mimetic compounds are useful for treating disorders characterised by low  
CC red blood cell levels such as anaemia. The TPO-mimetic comprising  
CC compounds are useful for treating conditions that involve an existing  
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic  
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,  
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35695 to ABB35777  
CC represent amino acid and nucleic acid sequences used in the  
CC exemplification of the present invention  
XX  
XX Sequence 20 AA;  
SQ

Query Match 3.9%; Score 7; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 LPLGGGG 170  
DB 13 LPLGGGG 19

RESULT 28  
ADC99215  
ID ADC99215 standard; peptide; 20 AA.  
XX  
XX  
AC ADC99215;  
XX  
XX  
DT 01-JAN-2004 (first entry)  
XX  
XX Cancer-related DGI-2-binder peptide - SEQ ID 48.  
DE  
XX  
XX cytostatic; cancer; gene therapy; DGI-2; DGI-5; DGI-7; DGI-9; Hras;  
KW leptin; VEGF; vascular endothelial growth factor receptor; VEGF-R1;  
KW VEGF-R2; VEGF-R3; FLT1; FMS-related tyrosine kinase 1; FLK1; KDR;  
KW Kinase insert domain protein receptor; EGFR; epidermal growth factor;  
KW FGFR1; fibroblast growth factor; Tie-1.  
XX  
XX Unidentified.  
OS  
XX  
XX WO2003035839-A2.  
PN  
XX  
XX 01-MAY-2003.  
PD  
XX  
XX 24-OCT-2002; 2002WO-US034021.  
PF

XX 24-OCT-2001; 2001US-0345471P.  
PR  
XX (DGI-) DGI BIOTECHNOLOGIES INC.  
PA  
XX  
PI Pillutia RC, Brissette R, Spruyt M, Dedova O, Blume A;  
XX Prendergast J, Goldstein N;  
XX WPI; 2003-457332/43.  
DR  
XX  
XX Selecting target and target binder pairs for preparing a composition for  
PT treating cancer by mixing in a reaction vessel phage expressing  
PT biological targets and phage expressing target binders.  
XX  
XX Claim 26; SEQ ID NO 48; 172pp; English.  
PS  
XX  
XX The invention relates to a novel method of selecting target and target  
CC binder pairs comprising mixing in a reaction vessel phage expressing  
CC biological targets and phage expressing target binders, each having  
CC distinguishable selection markers and selecting target and target binder  
CC pairs based on the selection markers. The molecules of the invention  
CC demonstrate cytostatic activity whilst the method may be useful for  
CC selecting target and target binder pairs for preparing a composition for  
CC treating cancer. Furthermore, the method may be utilised during gene  
CC therapy procedures. The current sequence is that of the cancer-related  
CC DGI-2-binder peptide of the invention.  
XX  
XX Sequence 20 AA;  
SQ

Query Match 3.9%; Score 7; DB 7; Length 20;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 166 LGGGGAR 172  
DB 11 LGGGGAR 17

RESULT 29  
AAB17944  
ID AAB17944 standard; peptide; 21 AA.  
XX  
XX AAB17944;  
AC  
XX 31-OCT-2000 (first entry)  
DT  
XX  
DE IL-1 R antagonist peptide sequence SEQ ID NO:1048.  
XX  
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;  
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;  
KW thrombosis; pharmaceutical.  
XX  
XX Synthetic.  
OS  
XX  
XX WO200024782-A2.  
PN  
XX  
XX 04-MAY-2000.  
PD  
XX  
XX 25-OCT-1999; 99WO-US025044.  
PF  
XX  
XX 23-OCT-1998; 98US-0105371P.  
PR  
XX 22-OCT-1999; 99US-00428082.  
PR  
XX (AMGE-) AMGEN INC.  
PA  
XX  
XX Feige U, Liu C, Cheetham J, Boone TC;  
PI  
XX WPI; 2000-350702/30.  
DR  
XX

PT Novel composition of matter comprising an Fc domain and pharmacologically  
 PT active peptides, useful for treating cancer and autoimmune diseases.

PS Disclosure; Page 564; 608pp; English.

XX The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-  
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,  
 CC P3, and P4 = are each independently sequences of pharmacologically active  
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,  
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1  
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,  
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host  
 CC cells from the present invention can be used for producing pharmaceutical  
 CC compositions. The compositions are useful for treating cancer, asthma,  
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than  
 CC a Fab domain) can provide a longer half-life or incorporate functions  
 CC such as Fc receptor binding, protein A binding, complement fixation, and  
 CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to  
 CC AAB18003 represent nucleotide and amino acid sequences used in the  
 CC exemplification of the present invention

XX Sequence 21 AA;

Query Match 3.9%; Score 7; DB 3; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 LPLGGGG 170  
 Db 13 LPLGGGG 19  
 |||||  
 |||||

RESULT 30

AAU62481

ID AAU62481 standard; protein; 56 AA.

XX AAU62481;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #23377.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59626.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.

XX Example 1; SEQ ID NO 23676; 1069pp; English.

PS Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 56 AA;

Query Match 3.9%; Score 7; DB 4; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 LGGGGAR 172  
 Db 19 LGGGGAR 25  
 |||||  
 |||||

RESULT 31

ABM59000

ID ABM59000 standard; protein; 56 AA.

XX ABM59000;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #23676.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
 KW immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978925.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallie-Douglass J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64555.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1; SEQ ID NO 23676; 1481pp; English.



CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to immunogenic fragments of P. acnes polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising P. acnes polypeptides, via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 56 AA;

Query Match 3.9%; Score 7; DB 6; Length 56;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGAR 172  
DB 19 LGGGGAR 25

RESULT 32  
AAU44577  
ID AAU44577 standard; protein; 64 AA.

AC AAU44577;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #5473.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypervitosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US012865.

PR 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59523.

XX Propionibacterium acnes polypeptides and nucleic acids useful for vaccinating against and diagnosing infections, especially useful for treating acne vulgaris.

PS Example 1; SEQ ID NO 5772; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 64 AA;

Query Match 3.9%; Score 7; DB 4; Length 64;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 LTVKEIK 146  
DB 54 LTVKEIK 60

RESULT 33

ABM41096  
ID ABM41096 standard; protein; 64 AA.

XX AC ABM41096;

XX 20-OCT-2003 (first entry)

DE Propionibacterium acnes predicted ORF-encoded polypeptide #5772.

KW Acne vulgaris; antisborrheic; dermatological; antibacterial; immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO2003033515-A1.

XX 24-APR-2003.

PF 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieue-Douglass J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64452.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.  
 XX Example 1; SEQ ID NO 5772; 1481bp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to immunogenic fragments of P. acnes polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 64 AA;

Query Match 3.9%; Score 7; DB 6; Length 64;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 140 LTVKEIK 146  
 |||||  
 Db 54 LTVKEIK 60

RESULT 34  
 AAW27922  
 ID AAW27922 standard; protein; 91 AA.  
 XX AAW27922;  
 XX  
 XX 20-AUG-1998 (first entry)  
 XX Staphylococcus aureus protein of unknown function.  
 XX  
 XX Staphylococcus aureus protein; ribozyme; antisense sequence; control;  
 KW Staphylococcal gene; regulatory element; bacterial gene expression;  
 KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;  
 XX toxic shock syndrome.  
 XX Staphylococcus aureus.  
 XX W09730070-A1.  
 XX  
 XX 21-AUG-1997.  
 XX  
 XX 19-FEB-1997; 97WO-US002318.  
 XX  
 XX 20-FEB-1996; 96US-0011888P.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;  
 PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;  
 XX

DR WPI; 1997-424969/39.  
 XX N-PSDB; AAT83884.  
 PT Novel polypeptide(s) from Staphylococcus aureus strain WCJH29 - used to  
 PT isolate antimicrobial compounds, and in vaccines against S. aureus  
 XX infection.  
 XX  
 PS Claim 6; Page 351-352; 989pp; English.  
 XX  
 CC The present sequence represents a Staphylococcus aureus protein of  
 CC unknown function. The DNA sequence was isolated from a library of clones  
 CC of S. aureus WCJH 29 in Escherichia coli. The DNA sequence can be used in  
 CC the construction of ribozymes and antisense sequences to control the  
 CC expression of Staphylococcal genes. The DNA sequence is also useful as a  
 CC source of regulatory elements for the control of bacterial gene  
 CC expression. The present protein may be used to produce vaccines to enable  
 CC a host to produce specific antibodies with antibacterial action. These  
 CC vaccines and antibodies would protect a host against invasion by S.  
 CC aureus, and conditions relating to Staphylococcal infection, e.g.  
 CC Staphylococcal food poisoning, scaled skin syndrome, and toxic shock  
 CC syndrome  
 XX  
 XX Sequence 91 AA;  
 CC  
 Query Match 3.9%; Score 7; DB 2; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 8 VTSIVTI 14  
 |||||  
 Db 44 VTSIVTI 50

RESULT 35  
 AAG35331  
 ID AAG35331 standard; protein; 93 AA.  
 XX AAG35331;  
 XX  
 XX 18-OCT-2000 (first entry)  
 XX Zea mays protein fragment SEQ ID NO: 43143.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.  
 XX  
 OS Zea mays subsp. mays.  
 XX  
 XX EPI033405-A2.  
 XX  
 XX 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-00301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825P.  
 XX 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-014091P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 15-JUL-1999; 99US-0143624P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0132485P.  
PR 22-JUL-1999; 99US-0132486P.  
PR 22-JUL-1999; 99US-0132487P.  
PR 23-JUL-1999; 99US-0132863P.  
PR 23-JUL-1999; 99US-0134256P.  
PR 23-JUL-1999; 99US-0134218P.  
PR 23-JUL-1999; 99US-0134219P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149226P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149823P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151068P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.

PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161040P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 25-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 3.9%; Score 7; DB 3; Length 93;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 166 LGGGGAR 172  
Db 13 LGGGGAR 19  
|||||  
|  
  
RESULT 36  
AAO11158  
ID AAO11158 standard; protein; 103 AA.  
AC AAO111158;  
XX XX  
DT 06-NOV-2001 (first entry)  
DE Human polypeptide SEQ ID NO 25050.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
XX WO200164835-A2.  
XX  
XX 07-SEP-2001.  
XX  
XX 26-FEB-2001; 2001WO-US004927.  
XX  
XX 28-FEB-2000; 2000US-00515126.  
PR 18-MAY-2000; 2000US-00577409.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-514838/56.  
XX N-PSDB; AAI91089.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
XX  
XX Claim 20; SEQ ID NO 25050; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93941) and  
CC the encoded proteins (AAO00010-AAO13910), that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 103 AA;  
  
Query Match 3.9%; Score 7; DB 4; Length 103;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 16 LLTSITS 22  
Db 85 LLTSITS 91  
|||||  
|  
  
RESULT 37  
ABP08245  
ID ABP08245 standard; protein; 118 AA.  
XX  
AC ABP08245;  
XX  
DT 24-JUN-2002 (first entry)  
DE Human ORFX protein sequence SEQ ID NO:16472.  
XX  
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myasthenia gravis.  
XX  
XX Homo sapiens.  
OS  
XX WO200192523-A2.  
XX  
XX 06-DEC-2001.  
XX  
XX 29-MAY-2001; 2001WO-US010836.  
XX  
XX 30-MAY-2000; 2000US-0206132P.  
PR 29-AUG-2000; 2000US-0228716P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shimkets RA, Leach MD;  
XX WPI; 2002-106308/14.  
XX N-PSDB; ABN23997.  
XX  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.  
XX  
XX Disclosure; SEQ ID NO 16472; 1037pp; English.  
XX  
XX The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 118 AA;  
 SQ

Query Match 3.9%; Score 7; DB 5; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GGGARIA 174  
 |||||  
 Db 105 GGGARIA 111

RESULT 38  
 ABU29948  
 ID ABU29948 standard; protein; 123 AA.  
 XX AC  
 XX ABU29948;  
 XX 19-JUN-2003 (first entry)  
 DT  
 XX Protein encoded by Prokaryotic essential gene #15475.  
 DE  
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 XX Enterococcus faecium.  
 OS  
 XX WO200277183-A2.  
 PN  
 XX 03-OCT-2002.  
 PD  
 XX 21-MAR-2002; 2002WO-US009107.  
 PF  
 XX 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 XX WPI: 2003-029926/02.  
 DR N-PSDB; ACA33618.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 25; SEQ ID NO 57872; 1766pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: the sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 123 AA;  
 SQ

Query Match 3.9%; Score 7; DB 6; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 TDLKGLP 67  
 |||||  
 Db 95 TDLKGLP 101

RESULT 39  
 ABM65961  
 ID ABM65961 standard; protein; 124 AA.  
 XX AC  
 XX ABM65961;  
 XX 20-OCT-2003 (first entry)  
 DT  
 XX Propionibacterium acnes immunogenic polypeptide #30637.  
 DE  
 XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
 KW immunostimulant; immune response; vaccine; immunogenic.  
 XX Propionibacterium acnes.  
 OS  
 XX WO2003033515-A1.  
 PN  
 XX 24-APR-2003.  
 PD  
 XX 11-OCT-2002; 2002WO-US032727.  
 PF  
 XX 15-OCT-2001; 2001US-00978825.  
 PR  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
 PI Zhang X, Wang S, Jen S, Iodes MJ, Benson DR, Jones R, Carter D;  
 PI Barth B, Valliee-Dougliss J;  
 XX  
 XX WPI: 2003-381789/36.  
 DR  
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a *P. acnes* protein.  
 XX  
 XX Claim 7; SEQ ID NO 30637; 148pp; English.  
 XX  
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 CC encoding a Propionibacterium acnes protein. The invention also relates to  
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
 CC immunogenic fragments of *P. acnes* polypeptides. The invention

CC additionally encompasses expression vectors and host cells comprising a  
 CC polynucleotide of the invention; antibodies against polypeptides of the  
 CC invention; fusion proteins comprising a polypeptide of the invention; a  
 CC method for stimulating an immune response specific for a P. acnes  
 CC polypeptide and an isolated T cell population comprising T cells prepared  
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
 CC antigen-presenting cells that express the polypeptide); a method and kit  
 CC for detecting or determining the presence or absence of P. acnes in a  
 CC patient; and a method for inhibiting the development of P. acnes in a  
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
 CC proteins, T cell populations or antigen-presenting cells that express the  
 CC polypeptides are useful for diagnosing, preventing or treating acne  
 CC vulgaris, or for stimulating an immune response specific for a P. acnes  
 CC protein. The polynucleotides can also be used as probes or primers for  
 CC nucleic acid hybridisation. The vaccine composition is useful for the  
 CC stimulation of an immune response against P. acnes, or for treating acne,  
 CC and the kit is useful for performing a diagnostic assay. The present  
 CC sequence represents a specifically claimed P. acnes polypeptide which is  
 CC thought to contain an immunogenic region. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 124 AA;

Query Match 3.9%; Score 7; DB 6; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 LPLGGGG 170  
 Db 10 LPLGGGG 16  
 |||||

RESULT 40  
 ABU29214  
 ID ABU29214 standard; protein; 124 AA.  
 XX  
 AC ABU29214;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #14741.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS Enterococcus faecalis.  
 XX  
 PN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 XX  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 FA (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA33084.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 57138; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: the sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 124 AA;

Query Match 3.9%; Score 7; DB 6; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 TDLKGLP 67  
 Db 95 TDLKGLP 101  
 |||||

Search completed: October 26, 2004, 10:00:19  
 Job time : 212 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 09:54:35 ; Search time 40 Seconds  
(without alignments)  
298.431 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 180

Sequence: 1 MKIKLFFVTISIVTISLLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/1aa/5A COMB.pep:\*
- 2: /cgn2\_6/prodata/1/1aa/5B COMB.pep:\*
- 3: /cgn2\_6/prodata/1/1aa/6A COMB.pep:\*
- 4: /cgn2\_6/prodata/1/1aa/6B COMB.pep:\*
- 5: /cgn2\_6/prodata/1/1aa/6CTUS COMB.pep:\*
- 6: /cgn2\_6/prodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 16    | 8.9         | 213    | 4  | US-09-328-352-6050   |
| 2          | 9     | 5.0         | 14     | 1  | US-07-641-346B-2     |
| 3          | 9     | 5.0         | 144    | 4  | US-09-489-039A-9511  |
| 4          | 9     | 5.0         | 154    | 1  | US-07-641-346B-1     |
| 5          | 8     | 4.4         | 759    | 1  | US-08-676-967-1      |
| 6          | 8     | 4.4         | 759    | 2  | US-08-676-974-1      |
| 7          | 8     | 4.4         | 759    | 2  | US-09-098-487-1      |
| 8          | 7     | 3.9         | 120    | 4  | US-09-428-082B-1048  |
| 9          | 7     | 3.9         | 133    | 4  | US-09-107-532A-6096  |
| 10         | 7     | 3.9         | 148    | 4  | US-09-270-767-37923  |
| 11         | 7     | 3.9         | 148    | 4  | US-09-270-767-53140  |
| 12         | 7     | 3.9         | 153    | 4  | US-09-252-991A-21201 |
| 13         | 7     | 3.9         | 157    | 4  | US-09-134-000C-3711  |
| 14         | 7     | 3.9         | 178    | 4  | US-09-543-681A-4466  |
| 15         | 7     | 3.9         | 191    | 4  | US-09-270-767-37351  |
| 16         | 7     | 3.9         | 191    | 4  | US-09-270-767-52568  |
| 17         | 7     | 3.9         | 248    | 4  | US-09-428-082B-1062  |
| 18         | 7     | 3.9         | 418    | 4  | US-09-252-991A-25726 |
| 19         | 7     | 3.9         | 500    | 4  | US-09-252-991A-25710 |
| 20         | 7     | 3.9         | 504    | 4  | US-09-252-991A-22291 |
| 21         | 7     | 3.9         | 535    | 4  | US-09-448-039A-11461 |
| 22         | 7     | 3.9         | 571    | 4  | US-09-248-796A-20375 |
| 23         | 7     | 3.9         | 605    | 4  | US-09-270-767-46196  |
| 24         | 7     | 3.9         | 674    | 4  | US-09-270-767-46168  |
| 25         | 7     | 3.9         | 679    | 4  | US-09-270-767-44535  |
| 26         | 7     | 3.9         | 814    | 4  | US-09-486-072-1      |
| 27         | 7     | 3.9         | 864    | 4  | US-09-604-978-11     |

|     |   |     |      |   |                      |                    |
|-----|---|-----|------|---|----------------------|--------------------|
| 28  | 7 | 3.9 | 864  | 4 | US-09-604-728-11     | Sequence 11, Appl  |
| 29  | 7 | 3.9 | 864  | 4 | US-10-325-878-11     | Sequence 11, Appl  |
| 30  | 7 | 3.9 | 1073 | 4 | US-09-180-245-2      | Sequence 2, Appl   |
| 31  | 7 | 3.9 | 1073 | 4 | US-09-819-249-2      | Sequence 2, Appl   |
| 32  | 7 | 3.9 | 1075 | 1 | US-07-623-033-2      | Sequence 2, Appl   |
| 33  | 6 | 3.3 | 14   | 6 | 5245013-14           | Patent No. 5245013 |
| 34  | 6 | 3.3 | 16   | 6 | 5245013-15           | Patent No. 5245013 |
| 35  | 6 | 3.3 | 18   | 3 | US-09-077-354B-5     | Sequence 5, Appl   |
| 36  | 6 | 3.3 | 20   | 6 | 5245013-11           | Patent No. 5245013 |
| 37  | 6 | 3.3 | 25   | 4 | US-09-383-062-54     | Sequence 54, Appl  |
| 38  | 6 | 3.3 | 25   | 4 | US-09-270-767-34704  | Sequence 34704, A  |
| 39  | 6 | 3.3 | 25   | 4 | US-09-270-767-49921  | Sequence 49921, A  |
| 40  | 6 | 3.3 | 32   | 2 | US-08-023-980B-27    | Sequence 27, Appl  |
| 41  | 6 | 3.3 | 32   | 2 | US-08-023-980B-29    | Sequence 29, Appl  |
| 42  | 6 | 3.3 | 32   | 2 | US-08-486-953A-22    | Sequence 22, Appl  |
| 43  | 6 | 3.3 | 32   | 2 | US-08-486-953A-24    | Sequence 24, Appl  |
| 44  | 6 | 3.3 | 32   | 4 | US-08-204-052-22     | Sequence 22, Appl  |
| 45  | 6 | 3.3 | 32   | 4 | US-08-204-052-24     | Sequence 24, Appl  |
| 46  | 6 | 3.3 | 33   | 2 | US-08-023-980B-32    | Sequence 32, Appl  |
| 47  | 6 | 3.3 | 33   | 2 | US-08-486-953A-27    | Sequence 27, Appl  |
| 48  | 6 | 3.3 | 33   | 4 | US-08-204-052-27     | Sequence 27, Appl  |
| 49  | 6 | 3.3 | 46   | 4 | US-09-270-767-34655  | Sequence 34655, A  |
| 50  | 6 | 3.3 | 46   | 4 | US-09-270-767-49872  | Sequence 49872, A  |
| 51  | 6 | 3.3 | 57   | 4 | US-09-270-767-40930  | Sequence 40930, A  |
| 52  | 6 | 3.3 | 57   | 4 | US-09-270-767-56146  | Sequence 56146, A  |
| 53  | 6 | 3.3 | 60   | 4 | US-09-328-352-5980   | Sequence 5980, Ap  |
| 54  | 6 | 3.3 | 65   | 4 | US-09-513-999C-7586  | Sequence 7586, Ap  |
| 55  | 6 | 3.3 | 67   | 4 | US-09-248-796A-24134 | Sequence 24134, A  |
| 56  | 6 | 3.3 | 72   | 4 | US-09-107-532A-7207  | Sequence 7207, Ap  |
| 57  | 6 | 3.3 | 75   | 4 | US-09-270-767-40037  | Sequence 40037, A  |
| 58  | 6 | 3.3 | 75   | 4 | US-09-270-767-55253  | Sequence 55253, A  |
| 59  | 6 | 3.3 | 76   | 3 | US-09-293-505-4      | Sequence 4, Appl   |
| 60  | 6 | 3.3 | 76   | 4 | US-09-060-939A-4     | Sequence 4, Appl   |
| 61  | 6 | 3.3 | 79   | 4 | US-09-621-976-7000   | Sequence 7000, Ap  |
| 62  | 6 | 3.3 | 99   | 4 | US-09-621-976-4323   | Sequence 4323, Ap  |
| 63  | 6 | 3.3 | 105  | 4 | US-09-107-532A-6573  | Sequence 6573, Ap  |
| 64  | 6 | 3.3 | 105  | 4 | US-09-543-681A-6080  | Sequence 6080, Ap  |
| 65  | 6 | 3.3 | 108  | 3 | US-09-134-001C-3086  | Sequence 3086, Ap  |
| 66  | 6 | 3.3 | 109  | 4 | US-09-107-532A-4999  | Sequence 4999, Ap  |
| 67  | 6 | 3.3 | 109  | 4 | US-09-107-532A-6406  | Sequence 6406, Ap  |
| 68  | 6 | 3.3 | 110  | 4 | US-09-328-352-7873   | Sequence 7873, Ap  |
| 69  | 6 | 3.3 | 111  | 4 | US-09-134-000C-4784  | Sequence 4784, Ap  |
| 70  | 6 | 3.3 | 120  | 4 | US-09-270-767-31866  | Sequence 31866, A  |
| 71  | 6 | 3.3 | 120  | 4 | US-09-270-767-47083  | Sequence 47083, A  |
| 72  | 6 | 3.3 | 120  | 4 | US-09-248-796A-22913 | Sequence 22913, A  |
| 73  | 6 | 3.3 | 121  | 2 | US-08-887-352B-2     | Sequence 2, Appl   |
| 74  | 6 | 3.3 | 121  | 3 | US-09-109-207C-2     | Sequence 2, Appl   |
| 75  | 6 | 3.3 | 121  | 3 | US-09-296-005-2      | Sequence 2, Appl   |
| 76  | 6 | 3.3 | 121  | 4 | US-09-920-171-2      | Sequence 2, Appl   |
| 77  | 6 | 3.3 | 121  | 4 | US-09-716-028-2      | Sequence 2, Appl   |
| 78  | 6 | 3.3 | 121  | 4 | US-10-113-996-2      | Sequence 2, Appl   |
| 79  | 6 | 3.3 | 123  | 4 | US-09-248-796A-21884 | Sequence 21884, A  |
| 80  | 6 | 3.3 | 125  | 4 | US-09-270-767-35574  | Sequence 35574, A  |
| 81  | 6 | 3.3 | 125  | 4 | US-09-270-767-50791  | Sequence 50791, A  |
| 82  | 6 | 3.3 | 132  | 4 | US-09-248-796A-25630 | Sequence 25630, A  |
| 83  | 6 | 3.3 | 134  | 3 | US-08-466-151-3      | Sequence 3, Appl   |
| 84  | 6 | 3.3 | 134  | 3 | US-08-466-163B-3     | Sequence 3, Appl   |
| 85  | 6 | 3.3 | 134  | 3 | US-09-186-276B-48    | Sequence 48, Appl  |
| 86  | 6 | 3.3 | 134  | 4 | US-08-842-445-48     | Sequence 48, Appl  |
| 87  | 6 | 3.3 | 134  | 4 | US-09-186-188B-48    | Sequence 48, Appl  |
| 88  | 6 | 3.3 | 134  | 4 | US-09-802-096-3      | Sequence 3, Appl   |
| 89  | 6 | 3.3 | 134  | 4 | US-09-802-077-3      | Sequence 3, Appl   |
| 90  | 6 | 3.3 | 134  | 4 | US-09-513-999C-7863  | Sequence 7863, Ap  |
| 91  | 6 | 3.3 | 139  | 3 | US-09-134-001C-4916  | Sequence 4916, Ap  |
| 92  | 6 | 3.3 | 139  | 4 | US-09-270-767-31757  | Sequence 31757, A  |
| 93  | 6 | 3.3 | 139  | 4 | US-09-270-767-46974  | Sequence 46974, A  |
| 94  | 6 | 3.3 | 142  | 4 | US-09-489-039A-8382  | Sequence 8382, Ap  |
| 95  | 6 | 3.3 | 145  | 3 | US-08-946-329A-55    | Sequence 55, Appl  |
| 96  | 6 | 3.3 | 149  | 3 | US-08-679-493A-200   | Sequence 200, App  |
| 97  | 6 | 3.3 | 149  | 4 | US-09-735-846-4      | Sequence 4, Appl   |
| 98  | 6 | 3.3 | 150  | 2 | US-08-722-050-11     | Sequence 11, Appl  |
| 99  | 6 | 3.3 | 150  | 3 | US-08-679-493A-199   | Sequence 199, App  |
| 100 | 6 | 3.3 | 150  | 4 | US-09-883-985-11     | Sequence 11, Appl  |

|     |   |     |     |   |                      |                   |     |   |     |     |   |                      |                    |
|-----|---|-----|-----|---|----------------------|-------------------|-----|---|-----|-----|---|----------------------|--------------------|
| 101 | 6 | 3.3 | 153 | 3 | US-08-679-493A-202   | Sequence 202, App | 174 | 6 | 3.3 | 254 | 4 | US-09-543-681A-5091  | Sequence 5091, Ap  |
| 102 | 6 | 3.3 | 154 | 3 | US-08-679-493A-211   | Sequence 211, App | 175 | 6 | 3.3 | 257 | 4 | US-09-248-796A-19035 | Sequence 19035, A  |
| 103 | 6 | 3.3 | 154 | 4 | US-09-489-039A-8261  | Sequence 8261, Ap | 176 | 6 | 3.3 | 260 | 4 | US-09-252-991A-21388 | Sequence 21388, A  |
| 104 | 6 | 3.3 | 154 | 4 | US-09-270-767-46484  | Sequence 46484, A | 177 | 6 | 3.3 | 262 | 4 | US-09-328-352-6552   | Sequence 6552, Ap  |
| 105 | 6 | 3.3 | 156 | 4 | US-09-489-039A-11508 | Sequence 11508, A | 178 | 6 | 3.3 | 262 | 4 | US-09-252-991A-30092 | Sequence 30092, A  |
| 106 | 6 | 3.3 | 156 | 4 | US-09-270-767-37406  | Sequence 37406, A | 179 | 6 | 3.3 | 264 | 1 | US-08-157-005-5      | Sequence 5, Appli  |
| 107 | 6 | 3.3 | 156 | 4 | US-09-270-767-52623  | Sequence 52623, A | 180 | 6 | 3.3 | 265 | 2 | US-08-799-464A-18    | Sequence 18, Appli |
| 108 | 6 | 3.3 | 157 | 4 | US-09-252-991A-23851 | Sequence 23851, A | 181 | 6 | 3.3 | 265 | 2 | US-08-747-863-5      | Sequence 5, Appli  |
| 109 | 6 | 3.3 | 163 | 1 | US-08-475-213-5      | Sequence 5, Appli | 182 | 6 | 3.3 | 265 | 3 | US-08-478-316-23     | Sequence 23, Appl  |
| 110 | 6 | 3.3 | 163 | 5 | PCT-US93-11703-1     | Sequence 1, Appli | 183 | 6 | 3.3 | 265 | 3 | US-09-019-793A-23    | Sequence 23, Appl  |
| 111 | 6 | 3.3 | 173 | 4 | US-09-248-796A-22963 | Sequence 22963, A | 184 | 6 | 3.3 | 265 | 4 | US-09-565-864-5      | Sequence 5, Appli  |
| 112 | 6 | 3.3 | 174 | 4 | US-09-252-991A-30989 | Sequence 30989, A | 185 | 6 | 3.3 | 265 | 4 | US-08-301-435-75     | Sequence 75, Appl  |
| 113 | 6 | 3.3 | 180 | 4 | US-09-615-192A-313   | Sequence 313, App | 186 | 6 | 3.3 | 265 | 4 | US-09-601-326-23     | Sequence 23, Appl  |
| 114 | 6 | 3.3 | 180 | 4 | US-09-270-767-46827  | Sequence 46827, A | 187 | 6 | 3.3 | 265 | 5 | PCT-US95-09927-18    | Sequence 18, Appl  |
| 115 | 6 | 3.3 | 186 | 4 | US-09-252-991A-27076 | Sequence 27076, A | 188 | 6 | 3.3 | 265 | 5 | PCT-US95-10904-75    | Sequence 75, Appl  |
| 116 | 6 | 3.3 | 186 | 4 | US-09-270-767-39985  | Sequence 39985, A | 189 | 6 | 3.3 | 266 | 4 | US-09-134-000C-6091  | Sequence 6091, Ap  |
| 117 | 6 | 3.3 | 190 | 4 | US-09-252-991A-18113 | Sequence 18113, A | 190 | 6 | 3.3 | 267 | 4 | US-09-252-991A-32647 | Sequence 32647, A  |
| 118 | 6 | 3.3 | 194 | 4 | US-09-270-767-45495  | Sequence 45495, A | 191 | 6 | 3.3 | 268 | 4 | US-09-583-110-4304   | Sequence 4304, Ap  |
| 119 | 6 | 3.3 | 195 | 3 | US-08-679-493A-189   | Sequence 189, App | 192 | 6 | 3.3 | 269 | 4 | US-09-071-252-18     | Sequence 18, Appl  |
| 120 | 6 | 3.3 | 196 | 3 | US-09-134-001C-4286  | Sequence 4286, Ap | 193 | 6 | 3.3 | 270 | 4 | US-09-107-532A-4153  | Sequence 4153, Ap  |
| 121 | 6 | 3.3 | 196 | 4 | US-09-540-236-3717   | Sequence 3717, Ap | 194 | 6 | 3.3 | 271 | 4 | US-09-809-665A-6     | Sequence 6, Appli  |
| 122 | 6 | 3.3 | 197 | 1 | US-08-186-811-4      | Sequence 4, Appli | 195 | 6 | 3.3 | 273 | 4 | US-09-252-991A-16693 | Sequence 16693, A  |
| 123 | 6 | 3.3 | 197 | 1 | US-08-261-660A-2     | Sequence 2, Appli | 196 | 6 | 3.3 | 273 | 4 | US-09-710-279-84     | Sequence 84, Appl  |
| 124 | 6 | 3.3 | 197 | 1 | US-08-261-660A-45    | Sequence 45, Appl | 197 | 6 | 3.3 | 274 | 3 | US-09-141-821-3      | Sequence 3, Appli  |
| 125 | 6 | 3.3 | 197 | 1 | US-08-261-660A-51    | Sequence 51, Appl | 198 | 6 | 3.3 | 275 | 3 | US-09-134-001C-3732  | Sequence 3732, Ap  |
| 126 | 6 | 3.3 | 197 | 1 | US-08-274-303-4      | Sequence 4, Appli | 199 | 6 | 3.3 | 286 | 4 | US-09-710-279-402    | Sequence 402, App  |
| 127 | 6 | 3.3 | 197 | 1 | US-08-377-391A-4     | Sequence 4, Appli | 200 | 6 | 3.3 | 288 | 3 | US-09-134-001C-4321  | Sequence 4321, Ap  |
| 128 | 6 | 3.3 | 197 | 2 | US-08-215-089-2      | Sequence 2, Appli | 201 | 6 | 3.3 | 289 | 4 | US-09-270-767-39365  | Sequence 39365, A  |
| 129 | 6 | 3.3 | 197 | 2 | US-08-779-400-4      | Sequence 4, Appli | 202 | 6 | 3.3 | 289 | 4 | US-09-270-767-54582  | Sequence 54582, A  |
| 130 | 6 | 3.3 | 197 | 3 | US-09-146-620-4      | Sequence 4, Appli | 203 | 6 | 3.3 | 290 | 4 | US-09-252-991A-21717 | Sequence 21717, A  |
| 131 | 6 | 3.3 | 197 | 3 | US-09-280-909A-2     | Sequence 2, Appli | 204 | 6 | 3.3 | 295 | 4 | US-09-248-796A-21485 | Sequence 21485, A  |
| 132 | 6 | 3.3 | 197 | 3 | US-09-280-909A-45    | Sequence 45, Appl | 205 | 6 | 3.3 | 296 | 3 | US-09-134-001C-4473  | Sequence 4473, Ap  |
| 133 | 6 | 3.3 | 197 | 4 | US-09-280-909A-51    | Sequence 51, Appl | 206 | 6 | 3.3 | 297 | 3 | US-09-632-947B-4     | Sequence 4, Appli  |
| 134 | 6 | 3.3 | 197 | 4 | US-09-486-147-40     | Sequence 40, Appl | 207 | 6 | 3.3 | 300 | 2 | US-08-897-340-32     | Sequence 32, Appl  |
| 135 | 6 | 3.3 | 197 | 5 | PCT-US94-06331-2     | Sequence 2, Appli | 208 | 6 | 3.3 | 300 | 3 | US-09-252-329-32     | Sequence 32, Appl  |
| 136 | 6 | 3.3 | 197 | 5 | PCT-US94-07834-4     | Sequence 4, Appli | 209 | 6 | 3.3 | 301 | 4 | US-09-328-352-7777   | Sequence 7777, Ap  |
| 137 | 6 | 3.3 | 197 | 5 | PCT-US95-03384-2     | Sequence 2, Appli | 210 | 6 | 3.3 | 302 | 4 | US-09-134-000C-5216  | Sequence 5216, Ap  |
| 138 | 6 | 3.3 | 198 | 1 | US-08-261-660A-52    | Sequence 52, Appl | 211 | 6 | 3.3 | 305 | 4 | US-09-252-991A-17702 | Sequence 17702, A  |
| 139 | 6 | 3.3 | 198 | 1 | US-08-261-660A-53    | Sequence 53, Appl | 212 | 6 | 3.3 | 306 | 4 | US-09-252-991A-18671 | Sequence 18671, A  |
| 140 | 6 | 3.3 | 198 | 3 | US-09-280-909A-52    | Sequence 52, Appl | 213 | 6 | 3.3 | 308 | 4 | US-09-252-991A-24816 | Sequence 24816, A  |
| 141 | 6 | 3.3 | 198 | 3 | US-09-280-909A-53    | Sequence 53, Appl | 214 | 6 | 3.3 | 310 | 4 | US-09-602-787A-662   | Sequence 662, App  |
| 142 | 6 | 3.3 | 199 | 1 | US-08-261-660A-46    | Sequence 46, Appl | 215 | 6 | 3.3 | 311 | 4 | US-09-270-767-42387  | Sequence 42387, A  |
| 143 | 6 | 3.3 | 199 | 1 | US-08-261-660A-48    | Sequence 48, Appl | 216 | 6 | 3.3 | 314 | 4 | US-09-489-039A-9630  | Sequence 9630, Ap  |
| 144 | 6 | 3.3 | 199 | 3 | US-09-280-909A-46    | Sequence 46, Appl | 217 | 6 | 3.3 | 315 | 3 | US-09-134-001C-3561  | Sequence 3561, Ap  |
| 145 | 6 | 3.3 | 199 | 3 | US-09-280-909A-48    | Sequence 48, Appl | 218 | 6 | 3.3 | 317 | 4 | US-09-252-991A-22776 | Sequence 22776, A  |
| 146 | 6 | 3.3 | 199 | 3 | US-09-351-229-2      | Sequence 2, Appli | 219 | 6 | 3.3 | 319 | 4 | US-09-328-352-4255   | Sequence 4255, Ap  |
| 147 | 6 | 3.3 | 199 | 4 | US-09-198-452A-595   | Sequence 595, App | 220 | 6 | 3.3 | 324 | 4 | US-09-489-039A-11252 | Sequence 11252, A  |
| 148 | 6 | 3.3 | 200 | 4 | US-09-248-796A-16251 | Sequence 16251, A | 221 | 6 | 3.3 | 327 | 3 | US-09-154-874-8      | Sequence 8, Appli  |
| 149 | 6 | 3.3 | 205 | 1 | US-08-277-231A-5     | Sequence 5, Appli | 222 | 6 | 3.3 | 327 | 4 | US-08-931-668-8      | Sequence 8, Appli  |
| 150 | 6 | 3.3 | 205 | 2 | US-08-477-326-8      | Sequence 8, Appli | 223 | 6 | 3.3 | 327 | 4 | US-09-468-175-8      | Sequence 8, Appli  |
| 151 | 6 | 3.3 | 205 | 2 | US-08-477-326-8      | Sequence 8, Appli | 224 | 6 | 3.3 | 328 | 3 | US-09-253-316-28     | Sequence 28, Appl  |
| 152 | 6 | 3.3 | 207 | 4 | US-09-543-681A-4234  | Sequence 4234, Ap | 225 | 6 | 3.3 | 328 | 4 | US-09-961-403-2      | Sequence 2, Appli  |
| 153 | 6 | 3.3 | 208 | 2 | US-08-038-364-2      | Sequence 2, Appli | 226 | 6 | 3.3 | 328 | 6 | 5212074-4            | Patent No. 5212074 |
| 154 | 6 | 3.3 | 208 | 3 | US-08-718-904-5      | Sequence 5, Appli | 227 | 6 | 3.3 | 329 | 4 | US-09-735-846-24     | Sequence 24, Appl  |
| 155 | 6 | 3.3 | 208 | 3 | US-09-181-974-2      | Sequence 2, Appli | 228 | 6 | 3.3 | 331 | 4 | US-09-735-846-23     | Sequence 23, Appl  |
| 156 | 6 | 3.3 | 208 | 3 | US-09-158-710-2      | Sequence 2, Appli | 229 | 6 | 3.3 | 332 | 4 | US-09-543-681A-4645  | Sequence 4645, Ap  |
| 157 | 6 | 3.3 | 208 | 3 | US-09-518-950-5      | Sequence 2, Appli | 230 | 6 | 3.3 | 333 | 3 | US-09-453-195A-2     | Sequence 2, Appli  |
| 158 | 6 | 3.3 | 208 | 4 | US-09-449-249-5      | Sequence 5, Appli | 231 | 6 | 3.3 | 341 | 4 | US-09-134-000C-4175  | Sequence 4175, Ap  |
| 159 | 6 | 3.3 | 208 | 4 | US-10-138-158-18     | Sequence 18, Appl | 232 | 6 | 3.3 | 342 | 4 | US-09-735-846-18     | Sequence 18, Appl  |
| 160 | 6 | 3.3 | 209 | 4 | US-09-328-352-5155   | Sequence 5155, Ap | 233 | 6 | 3.3 | 344 | 1 | US-08-891-254-7      | Sequence 7, Appli  |
| 161 | 6 | 3.3 | 209 | 4 | US-09-489-039A-10471 | Sequence 10471, A | 234 | 6 | 3.3 | 344 | 2 | US-08-819-539-7      | Sequence 7, Appli  |
| 162 | 6 | 3.3 | 221 | 4 | US-09-303-040-8      | Sequence 8, Appli | 235 | 6 | 3.3 | 344 | 2 | US-09-030-270A-7     | Sequence 7, Appli  |
| 163 | 6 | 3.3 | 222 | 3 | US-08-956-182-17     | Sequence 17, Appl | 236 | 6 | 3.3 | 344 | 3 | US-08-984-207-7      | Sequence 7, Appli  |
| 164 | 6 | 3.3 | 229 | 4 | US-09-489-039A-12532 | Sequence 12532, A | 237 | 6 | 3.3 | 344 | 3 | US-09-013-587-7      | Sequence 7, Appli  |
| 165 | 6 | 3.3 | 233 | 3 | US-09-069-023-38     | Sequence 38, Appl | 238 | 6 | 3.3 | 344 | 4 | US-09-086-118-27     | Sequence 27, Appl  |
| 166 | 6 | 3.3 | 234 | 4 | US-09-710-279-1380   | Sequence 1380, Ap | 239 | 6 | 3.3 | 344 | 4 | US-09-431-614-15     | Sequence 15, Appl  |
| 167 | 6 | 3.3 | 238 | 4 | US-09-328-352-6651   | Sequence 6651, Ap | 240 | 6 | 3.3 | 344 | 4 | US-09-735-846-22     | Sequence 22, Appl  |
| 168 | 6 | 3.3 | 240 | 4 | US-09-636-735A-2     | Sequence 2, Appli | 241 | 6 | 3.3 | 344 | 4 | US-09-809-665A-101   | Sequence 101, App  |
| 169 | 6 | 3.3 | 240 | 4 | US-09-636-735A-12    | Sequence 12, Appl | 242 | 6 | 3.3 | 344 | 5 | PCT-US96-08819-7     | Sequence 7, Appli  |
| 170 | 6 | 3.3 | 241 | 4 | US-09-153-599A-5     | Sequence 5, Appli | 243 | 6 | 3.3 | 349 | 4 | US-09-735-846-2      | Sequence 2, Appli  |
| 171 | 6 | 3.3 | 247 | 4 | US-09-134-000C-5222  | Sequence 5222, Ap | 244 | 6 | 3.3 | 349 | 4 | US-09-735-846-12     | Sequence 12, Appl  |
| 172 | 6 | 3.3 | 248 | 4 | US-09-248-796A-16541 | Sequence 16541, A | 245 | 6 | 3.3 | 351 | 4 | US-09-431-976-2      | Sequence 2, Appli  |
| 173 | 6 | 3.3 | 252 | 4 | US-09-198-452A-975   | Sequence 975, App | 246 | 6 | 3.3 | 351 | 4 | US-09-252-991A-25097 | Sequence 25097, A  |



|     |   |     |     |   |                       |                    |     |   |     |     |   |                      |                    |
|-----|---|-----|-----|---|-----------------------|--------------------|-----|---|-----|-----|---|----------------------|--------------------|
| 247 | 6 | 3.3 | 351 | 4 | US-09-021-286-2       | Sequence 2, Appli  | 320 | 6 | 3.3 | 461 | 3 | US-09-042-785A-7     | Sequence 7, Appli  |
| 248 | 6 | 3.3 | 352 | 4 | US-09-248-796A-26579  | Sequence 26579, A  | 321 | 6 | 3.3 | 461 | 3 | US-08-477-347-3      | Sequence 3, Appli  |
| 249 | 6 | 3.3 | 353 | 4 | US-09-252-991A-21002  | Sequence 21002, A  | 322 | 6 | 3.3 | 461 | 3 | US-09-006-353A-4     | Sequence 4, Appli  |
| 250 | 6 | 3.3 | 354 | 4 | US-09-252-991A-31625  | Sequence 31625, A  | 323 | 6 | 3.3 | 461 | 3 | US-08-476-862-2      | Sequence 2, Appli  |
| 251 | 6 | 3.3 | 355 | 4 | US-09-583-110-3998    | Sequence 3998, Ap  | 324 | 6 | 3.3 | 461 | 4 | US-09-573-986-4      | Sequence 4, Appli  |
| 252 | 6 | 3.3 | 356 | 4 | US-09-328-352-6575    | Sequence 6575, Ap  | 325 | 6 | 3.3 | 461 | 4 | US-08-406-824A-2     | Sequence 2, Appli  |
| 253 | 6 | 3.3 | 357 | 4 | US-10-140-002-106     | Sequence 106, App  | 326 | 6 | 3.3 | 461 | 4 | US-09-800-909-2      | Sequence 2, Appli  |
| 254 | 6 | 3.3 | 374 | 4 | US-09-252-991A-31294  | Sequence 31294, A  | 327 | 6 | 3.3 | 461 | 4 | US-09-758-124-2      | Sequence 2, Appli  |
| 255 | 6 | 3.3 | 376 | 4 | US-09-270-767-45545   | Sequence 45545, A  | 328 | 6 | 3.3 | 461 | 4 | US-09-800-908-3      | Sequence 3, Appli  |
| 256 | 6 | 3.3 | 377 | 4 | US-07-857-224B-88     | Sequence 88, Appli | 329 | 6 | 3.3 | 461 | 4 | US-09-896-096A-17    | Sequence 17, Appli |
| 257 | 6 | 3.3 | 377 | 4 | US-10-179-784-1       | Sequence 1, Appli  | 330 | 6 | 3.3 | 461 | 6 | 5395760-2            | Patent No. 5395760 |
| 258 | 6 | 3.3 | 380 | 5 | PCT-US91-02560-4      | Sequence 4, Appli  | 331 | 6 | 3.3 | 463 | 4 | US-09-252-991A-26534 | Sequence 26534, A  |
| 259 | 6 | 3.3 | 381 | 4 | PCT-US91-02591A-27174 | Sequence 27174, A  | 332 | 6 | 3.3 | 464 | 2 | US-08-477-451-18     | Sequence 18, Appli |
| 260 | 6 | 3.3 | 383 | 4 | US-09-266-965-100     | Sequence 100, App  | 333 | 6 | 3.3 | 470 | 4 | US-09-252-991A-29251 | Sequence 29251, A  |
| 261 | 6 | 3.3 | 384 | 4 | US-09-107-532A-5318   | Sequence 5318, Ap  | 334 | 6 | 3.3 | 471 | 4 | US-09-513-007-2      | Sequence 2, Appli  |
| 262 | 6 | 3.3 | 392 | 4 | US-09-252-991A-20001  | Sequence 20001, A  | 335 | 6 | 3.3 | 472 | 4 | US-09-520-781-22     | Sequence 22, Appli |
| 263 | 6 | 3.3 | 393 | 4 | US-09-600-099-6       | Sequence 6, Appli  | 336 | 6 | 3.3 | 473 | 4 | US-09-934-899-4      | Sequence 4, Appli  |
| 264 | 6 | 3.3 | 393 | 4 | US-09-328-352-7625    | Sequence 7625, Ap  | 337 | 6 | 3.3 | 473 | 4 | US-09-934-868-24     | Sequence 24, Appli |
| 265 | 6 | 3.3 | 394 | 3 | US-08-673-814-6       | Sequence 6, Appli  | 338 | 6 | 3.3 | 474 | 2 | US-08-650-000-4      | Sequence 4, Appli  |
| 266 | 6 | 3.3 | 394 | 3 | US-09-115-824-6       | Sequence 6, Appli  | 339 | 6 | 3.3 | 474 | 3 | US-09-042-785A-8     | Sequence 8, Appli  |
| 267 | 6 | 3.3 | 396 | 4 | US-09-252-991A-32927  | Sequence 32927, A  | 340 | 6 | 3.3 | 474 | 4 | US-09-758-124-4      | Sequence 4, Appli  |
| 268 | 6 | 3.3 | 400 | 1 | US-08-351-149-6       | Sequence 6, Appli  | 341 | 6 | 3.3 | 474 | 6 | 5395760-4            | Patent No. 5395760 |
| 269 | 6 | 3.3 | 400 | 1 | US-08-384-828-6       | Sequence 6, Appli  | 342 | 6 | 3.3 | 477 | 1 | US-07-847-562-2      | Sequence 2, Appli  |
| 270 | 6 | 3.3 | 400 | 3 | US-08-895-474-6       | Sequence 6, Appli  | 343 | 6 | 3.3 | 477 | 1 | US-08-240-328-2      | Sequence 2, Appli  |
| 271 | 6 | 3.3 | 402 | 3 | US-09-319-892-4       | Sequence 4, Appli  | 344 | 6 | 3.3 | 477 | 2 | US-08-990-849-2      | Sequence 2, Appli  |
| 272 | 6 | 3.3 | 404 | 4 | US-09-727-238-4       | Sequence 4, Appli  | 345 | 6 | 3.3 | 478 | 6 | 5245013-3            | Patent No. 5245013 |
| 273 | 6 | 3.3 | 404 | 4 | US-09-644-460-29      | Sequence 29, Appli | 346 | 6 | 3.3 | 478 | 4 | US-09-252-991A-31246 | Sequence 31246, A  |
| 274 | 6 | 3.3 | 405 | 4 | US-09-107-532A-6766   | Sequence 6766, Ap  | 347 | 6 | 3.3 | 479 | 3 | US-08-431-517F-17    | Sequence 17, Appli |
| 275 | 6 | 3.3 | 407 | 4 | US-09-328-352-7393    | Sequence 7393, Ap  | 348 | 6 | 3.3 | 479 | 4 | US-09-252-991A-25982 | Sequence 25982, A  |
| 276 | 6 | 3.3 | 409 | 4 | US-09-328-352-7559    | Sequence 7559, Ap  | 349 | 6 | 3.3 | 481 | 1 | US-08-186-811-2      | Sequence 2, Appli  |
| 277 | 6 | 3.3 | 410 | 4 | US-09-252-991A-25812  | Sequence 25812, A  | 350 | 6 | 3.3 | 481 | 1 | US-08-311-611A-98    | Sequence 98, Appli |
| 278 | 6 | 3.3 | 411 | 1 | US-07-684-135A-2      | Sequence 2, Appli  | 351 | 6 | 3.3 | 481 | 1 | US-08-372-783-98     | Sequence 98, Appli |
| 279 | 6 | 3.3 | 411 | 4 | US-09-270-767-58315   | Sequence 58315, A  | 352 | 6 | 3.3 | 481 | 1 | US-08-372-105-98     | Sequence 98, Appli |
| 280 | 6 | 3.3 | 412 | 4 | US-09-252-991A-19536  | Sequence 19536, A  | 353 | 6 | 3.3 | 481 | 1 | US-08-306-473A-98    | Sequence 98, Appli |
| 281 | 6 | 3.3 | 412 | 4 | US-09-252-991A-27252  | Sequence 27252, A  | 354 | 6 | 3.3 | 481 | 1 | US-08-261-660A-4     | Sequence 4, Appli  |
| 282 | 6 | 3.3 | 413 | 2 | US-08-846-762-15      | Sequence 15, Appli | 355 | 6 | 3.3 | 481 | 1 | US-08-209-762-98     | Sequence 98, Appli |
| 283 | 6 | 3.3 | 413 | 2 | US-08-846-762-88      | Sequence 88, Appli | 356 | 6 | 3.3 | 481 | 1 | US-08-473-344-98     | Sequence 98, Appli |
| 284 | 6 | 3.3 | 413 | 4 | US-09-543-681A-5108   | Sequence 5108, Ap  | 357 | 6 | 3.3 | 481 | 1 | US-08-274-303-6      | Sequence 6, Appli  |
| 285 | 6 | 3.3 | 413 | 4 | US-09-489-039A-12692  | Sequence 12692, A  | 358 | 6 | 3.3 | 481 | 2 | US-08-377-391A-2     | Sequence 2, Appli  |
| 286 | 6 | 3.3 | 416 | 4 | US-09-347-798-4       | Sequence 4, Appli  | 359 | 6 | 3.3 | 481 | 2 | US-08-215-089-9      | Sequence 9, Appli  |
| 287 | 6 | 3.3 | 418 | 4 | US-09-489-039A-14152  | Sequence 14152, A  | 360 | 6 | 3.3 | 481 | 2 | US-08-485-445A-98    | Sequence 98, Appli |
| 288 | 6 | 3.3 | 419 | 4 | US-09-710-279-2516    | Sequence 2516, Ap  | 361 | 6 | 3.3 | 481 | 2 | US-08-779-400-2      | Sequence 2, Appli  |
| 289 | 6 | 3.3 | 420 | 4 | US-09-489-039A-7218   | Sequence 7218, Ap  | 362 | 6 | 3.3 | 481 | 2 | US-08-955-660-2      | Sequence 2, Appli  |
| 290 | 6 | 3.3 | 420 | 4 | US-09-489-039A-8902   | Sequence 8902, Ap  | 363 | 6 | 3.3 | 481 | 3 | US-09-119-263-98     | Sequence 98, Appli |
| 291 | 6 | 3.3 | 420 | 4 | US-09-583-110-5043    | Sequence 5043, Ap  | 364 | 6 | 3.3 | 481 | 3 | US-08-657-162-98     | Sequence 98, Appli |
| 292 | 6 | 3.3 | 424 | 4 | US-09-107-532A-6341   | Sequence 6341, Ap  | 365 | 6 | 3.3 | 481 | 3 | US-09-224-480-98     | Sequence 98, Appli |
| 293 | 6 | 3.3 | 430 | 3 | US-09-256-000-23      | Sequence 23, Appli | 366 | 6 | 3.3 | 481 | 3 | US-09-093-539-98     | Sequence 98, Appli |
| 294 | 6 | 3.3 | 430 | 4 | US-09-917-254-90      | Sequence 90, Appli | 367 | 6 | 3.3 | 481 | 3 | US-08-431-517F-2     | Sequence 2, Appli  |
| 295 | 6 | 3.3 | 430 | 4 | US-10-034-015A-23     | Sequence 23, Appli | 368 | 6 | 3.3 | 481 | 3 | US-08-431-517F-7     | Sequence 7, Appli  |
| 296 | 6 | 3.3 | 433 | 4 | US-09-710-279-1332    | Sequence 1332, Ap  | 369 | 6 | 3.3 | 481 | 3 | US-08-431-517F-11    | Sequence 11, Appli |
| 297 | 6 | 3.3 | 435 | 3 | US-09-134-001C-4622   | Sequence 4622, Ap  | 370 | 6 | 3.3 | 481 | 3 | US-09-146-620-2      | Sequence 2, Appli  |
| 298 | 6 | 3.3 | 436 | 4 | US-09-543-681A-4395   | Sequence 4395, Ap  | 371 | 6 | 3.3 | 481 | 3 | US-09-395-453-2      | Sequence 2, Appli  |
| 299 | 6 | 3.3 | 437 | 3 | US-09-134-001C-3418   | Sequence 3418, Ap  | 372 | 6 | 3.3 | 481 | 3 | US-09-280-909A-4     | Sequence 4, Appli  |
| 300 | 6 | 3.3 | 438 | 4 | US-09-248-796A-23919  | Sequence 23919, A  | 373 | 6 | 3.3 | 481 | 3 | US-09-790-230-98     | Sequence 98, Appli |
| 301 | 6 | 3.3 | 448 | 4 | US-09-252-991A-30230  | Sequence 30230, A  | 374 | 6 | 3.3 | 481 | 4 | US-09-578-063-39     | Sequence 39, Appli |
| 302 | 6 | 3.3 | 449 | 4 | US-09-252-991A-31260  | Sequence 31260, A  | 375 | 6 | 3.3 | 481 | 5 | PCT-US94-02465-98    | Sequence 98, Appli |
| 303 | 6 | 3.3 | 452 | 1 | US-08-205-719-3       | Sequence 3, Appli  | 376 | 6 | 3.3 | 481 | 5 | PCT-US94-06931-4     | Sequence 4, Appli  |
| 304 | 6 | 3.3 | 452 | 4 | US-09-252-991A-25818  | Sequence 25818, A  | 377 | 6 | 3.3 | 481 | 5 | PCT-US95-00498-98    | Sequence 98, Appli |
| 305 | 6 | 3.3 | 452 | 4 | US-09-270-767-60765   | Sequence 60765, A  | 378 | 6 | 3.3 | 481 | 5 | PCT-US95-00656-98    | Sequence 98, Appli |
| 306 | 6 | 3.3 | 453 | 4 | US-09-252-991A-20755  | Sequence 20755, A  | 379 | 6 | 3.3 | 481 | 5 | PCT-US95-03384-9     | Sequence 9, Appli  |
| 307 | 6 | 3.3 | 453 | 4 | US-09-710-279-1516    | Sequence 1516, Ap  | 380 | 6 | 3.3 | 481 | 5 | PCT-US95-03384-9     | Sequence 9, Appli  |
| 308 | 6 | 3.3 | 454 | 1 | US-07-915-720D-17     | Sequence 17, Appli | 381 | 6 | 3.3 | 482 | 4 | US-08-311-731A-204   | Sequence 204, App  |
| 309 | 6 | 3.3 | 454 | 3 | US-09-025-543-17      | Sequence 17, Appli | 382 | 6 | 3.3 | 482 | 4 | US-09-248-796A-16662 | Sequence 16662, A  |
| 310 | 6 | 3.3 | 454 | 3 | US-09-134-001C-3501   | Sequence 3501, Ap  | 383 | 6 | 3.3 | 491 | 2 | US-08-942-819-2      | Sequence 2, Appli  |
| 311 | 6 | 3.3 | 456 | 3 | US-08-879-565-12      | Sequence 12, Appli | 384 | 6 | 3.3 | 491 | 4 | US-09-522-955A-2     | Sequence 2, Appli  |
| 312 | 6 | 3.3 | 456 | 3 | US-08-431-517F-8      | Sequence 8, Appli  | 385 | 6 | 3.3 | 491 | 4 | US-09-520-781-24     | Sequence 24, Appli |
| 313 | 6 | 3.3 | 456 | 3 | US-08-431-517F-12     | Sequence 12, Appli | 386 | 6 | 3.3 | 492 | 4 | US-09-792-024-91     | Sequence 91, Appli |
| 314 | 6 | 3.3 | 456 | 4 | US-09-252-991A-19417  | Sequence 19417, A  | 387 | 6 | 3.3 | 493 | 1 | US-08-362-512A-4     | Sequence 4, Appli  |
| 315 | 6 | 3.3 | 459 | 3 | US-09-538-414-2       | Sequence 2, Appli  | 388 | 6 | 3.3 | 493 | 3 | US-08-964-939-4      | Sequence 4, Appli  |
| 316 | 6 | 3.3 | 459 | 4 | US-09-252-991A-27280  | Sequence 27280, A  | 389 | 6 | 3.3 | 494 | 4 | US-09-489-039A-10060 | Sequence 10060, A  |
| 317 | 6 | 3.3 | 459 | 4 | US-10-074-273-2       | Sequence 2, Appli  | 390 | 6 | 3.3 | 503 | 3 | US-08-999-774A-10    | Sequence 10, Appli |
| 318 | 6 | 3.3 | 461 | 1 | US-08-385-229-2       | Sequence 2, Appli  | 391 | 6 | 3.3 | 503 | 4 | US-09-252-991A-26258 | Sequence 26258, A  |
| 319 | 6 | 3.3 | 461 | 2 | US-08-650-000-2       | Sequence 2, Appli  | 392 | 6 | 3.3 | 510 | 3 | US-09-134-001C-3368  | Sequence 3368, Ap  |

|     |   |     |     |   |                      |                   |     |   |     |      |   |                      |                    |
|-----|---|-----|-----|---|----------------------|-------------------|-----|---|-----|------|---|----------------------|--------------------|
| 393 | 6 | 3.3 | 511 | 1 | US-08-480-604A-20    | Sequence 20, Appl | 466 | 6 | 3.3 | 694  | 5 | PCT-US95-11856-2     | Sequence 2, Appl   |
| 394 | 6 | 3.3 | 511 | 2 | US-08-405-496A-20    | Sequence 20, Appl | 467 | 6 | 3.3 | 694  | 5 | PCT-US95-11878-2     | Sequence 2, Appl   |
| 395 | 6 | 3.3 | 511 | 3 | US-08-915-136-20     | Sequence 20, Appl | 468 | 6 | 3.3 | 700  | 4 | US-09-252-991A-21829 | Sequence 21829, A  |
| 396 | 6 | 3.3 | 511 | 3 | US-08-957-310-20     | Sequence 20, Appl | 469 | 6 | 3.3 | 707  | 4 | US-09-919-039-278    | Sequence 278, App  |
| 397 | 6 | 3.3 | 511 | 3 | US-08-957-310-20     | Sequence 20, Appl | 470 | 6 | 3.3 | 707  | 4 | US-09-538-092-993    | Sequence 993, App  |
| 398 | 6 | 3.3 | 511 | 4 | US-10-011-366-20     | Sequence 20, Appl | 471 | 6 | 3.3 | 711  | 4 | US-10-158-847-138    | Sequence 138, App  |
| 399 | 6 | 3.3 | 511 | 4 | US-09-084-517-20     | Sequence 20, Appl | 472 | 6 | 3.3 | 719  | 2 | US-09-003-217-2      | Sequence 2, Appl   |
| 400 | 6 | 3.3 | 523 | 4 | US-09-134-000C-14499 | Sequence 14499, A | 473 | 6 | 3.3 | 719  | 2 | US-09-003-217-2      | Sequence 2, Appl   |
| 401 | 6 | 3.3 | 523 | 4 | US-09-248-796A-14499 | Sequence 6006, Ap | 474 | 6 | 3.3 | 719  | 3 | US-08-286-870A-8     | Sequence 8, Appl   |
| 402 | 6 | 3.3 | 525 | 3 | US-08-984-618-17     | Sequence 17, Appl | 475 | 6 | 3.3 | 719  | 3 | US-08-975-762-59     | Sequence 59, Appl  |
| 403 | 6 | 3.3 | 525 | 3 | US-08-984-618-17     | Sequence 17, Appl | 476 | 6 | 3.3 | 719  | 3 | US-09-218-942-2      | Sequence 2, Appl   |
| 404 | 6 | 3.3 | 529 | 1 | US-08-178-477B-32    | Sequence 32, Appl | 477 | 6 | 3.3 | 719  | 3 | US-09-295-028-59     | Sequence 59, Appl  |
| 405 | 6 | 3.3 | 529 | 3 | US-09-304-121-2      | Sequence 20, Appl | 478 | 6 | 3.3 | 719  | 3 | US-09-106-582-59     | Sequence 59, Appl  |
| 406 | 6 | 3.3 | 533 | 3 | US-09-256-000-19     | Sequence 19, Appl | 479 | 6 | 3.3 | 719  | 3 | US-09-159-469-59     | Sequence 59, Appl  |
| 407 | 6 | 3.3 | 533 | 4 | US-10-034-015A-19    | Sequence 19, Appl | 480 | 6 | 3.3 | 728  | 4 | US-09-693-542-59     | Sequence 59, Appl  |
| 408 | 6 | 3.3 | 535 | 3 | US-08-286-870A-6     | Sequence 6, Appl  | 481 | 6 | 3.3 | 743  | 3 | US-09-489-039A-11954 | Sequence 11954, A  |
| 409 | 6 | 3.3 | 535 | 4 | US-09-270-767-42002  | Sequence 42002, A | 482 | 6 | 3.3 | 755  | 4 | US-09-077-354B-2     | Sequence 2, Appl   |
| 410 | 6 | 3.3 | 539 | 4 | US-09-252-991A-28826 | Sequence 28826, A | 483 | 6 | 3.3 | 757  | 4 | US-09-270-767-45272  | Sequence 45272, A  |
| 411 | 6 | 3.3 | 544 | 4 | US-09-248-796A-27039 | Sequence 27039, A | 484 | 6 | 3.3 | 761  | 4 | US-09-963-791-24     | Sequence 24, Appl  |
| 412 | 6 | 3.3 | 550 | 2 | US-08-816-155B-44    | Sequence 44, Appl | 485 | 6 | 3.3 | 761  | 4 | US-09-328-352-5650   | Sequence 5650, Ap  |
| 413 | 6 | 3.3 | 550 | 3 | US-09-079-587-44     | Sequence 44, Appl | 486 | 6 | 3.3 | 779  | 1 | US-08-375-134-12     | Sequence 12, Appl  |
| 414 | 6 | 3.3 | 551 | 4 | US-09-328-352-5333   | Sequence 5333, Ap | 487 | 6 | 3.3 | 779  | 5 | PCT-US95-15263-12    | Sequence 12, Appl  |
| 415 | 6 | 3.3 | 551 | 4 | US-09-252-991A-20358 | Sequence 20358, A | 488 | 6 | 3.3 | 783  | 4 | US-09-513-783A-176   | Sequence 176, App  |
| 416 | 6 | 3.3 | 552 | 4 | US-09-969-532-8      | Sequence 8, Appl  | 489 | 6 | 3.3 | 802  | 4 | US-09-823-240A-2     | Sequence 2, Appl   |
| 417 | 6 | 3.3 | 552 | 4 | US-09-792-024-84     | Sequence 84, Appl | 490 | 6 | 3.3 | 805  | 3 | US-08-989-299-2      | Sequence 2, Appl   |
| 418 | 6 | 3.3 | 555 | 4 | US-10-140-002-72     | Sequence 72, Appl | 491 | 6 | 3.3 | 805  | 4 | US-10-158-847-142    | Sequence 142, App  |
| 419 | 6 | 3.3 | 557 | 4 | US-09-252-991A-28083 | Sequence 28083, A | 492 | 6 | 3.3 | 805  | 4 | US-09-407-427-2      | Sequence 2, Appl   |
| 420 | 6 | 3.3 | 563 | 4 | US-09-969-532-6      | Sequence 6, Appl  | 493 | 6 | 3.3 | 823  | 4 | US-09-252-991A-27614 | Sequence 27614, A  |
| 421 | 6 | 3.3 | 565 | 4 | US-09-248-796A-20098 | Sequence 20098, A | 494 | 6 | 3.3 | 830  | 4 | US-09-252-991A-20619 | Sequence 20619, A  |
| 422 | 6 | 3.3 | 566 | 4 | US-09-969-532-4      | Sequence 4, Appl  | 495 | 6 | 3.3 | 830  | 4 | US-09-248-796A-18006 | Sequence 18006, A  |
| 423 | 6 | 3.3 | 575 | 4 | US-09-252-991A-19291 | Sequence 19291, A | 496 | 6 | 3.3 | 839  | 4 | US-09-489-039A-13252 | Sequence 13252, A  |
| 424 | 6 | 3.3 | 576 | 4 | US-09-489-039A-7859  | Sequence 7859, Ap | 497 | 6 | 3.3 | 843  | 4 | US-09-252-991A-32609 | Sequence 32609, A  |
| 425 | 6 | 3.3 | 577 | 4 | US-09-969-532-2      | Sequence 2, Appl  | 498 | 6 | 3.3 | 850  | 4 | US-09-583-110-4394   | Sequence 4394, Ap  |
| 426 | 6 | 3.3 | 580 | 3 | US-09-256-000-21     | Sequence 21, Appl | 499 | 6 | 3.3 | 850  | 4 | US-08-893-525-42     | Sequence 42, Appl  |
| 427 | 6 | 3.3 | 580 | 4 | US-10-034-015A-21    | Sequence 21, Appl | 500 | 6 | 3.3 | 856  | 4 | US-09-248-796A-16089 | Sequence 16089, A  |
| 428 | 6 | 3.3 | 581 | 4 | US-09-538-092-40     | Sequence 40, Appl | 501 | 6 | 3.3 | 859  | 3 | US-09-369-364A-5     | Sequence 5, Appl   |
| 429 | 6 | 3.3 | 581 | 4 | US-09-107-532A-5385  | Sequence 5385, Ap | 502 | 6 | 3.3 | 883  | 4 | US-09-248-796A-14418 | Sequence 14418, A  |
| 430 | 6 | 3.3 | 593 | 2 | US-08-987-466-1      | Sequence 1, Appl  | 503 | 6 | 3.3 | 885  | 1 | US-08-484-105-14     | Sequence 14, Appl  |
| 431 | 6 | 3.3 | 593 | 3 | US-09-240-359-1      | Sequence 1, Appl  | 504 | 6 | 3.3 | 885  | 1 | US-08-484-105-14     | Sequence 14, Appl  |
| 432 | 6 | 3.3 | 603 | 4 | US-09-270-767-46486  | Sequence 46486, A | 505 | 6 | 3.3 | 885  | 4 | US-09-543-681A-4596  | Sequence 4596, Ap  |
| 433 | 6 | 3.3 | 606 | 4 | US-09-538-092-798    | Sequence 798, App | 506 | 6 | 3.3 | 886  | 4 | US-09-969-532-16     | Sequence 16, Appl  |
| 434 | 6 | 3.3 | 608 | 1 | US-08-480-604A-21    | Sequence 21, Appl | 507 | 6 | 3.3 | 897  | 4 | US-09-969-532-14     | Sequence 14, Appl  |
| 435 | 6 | 3.3 | 608 | 2 | US-08-405-496A-21    | Sequence 21, Appl | 508 | 6 | 3.3 | 900  | 4 | US-09-969-532-12     | Sequence 12, Appl  |
| 436 | 6 | 3.3 | 608 | 3 | US-08-915-136-21     | Sequence 21, Appl | 509 | 6 | 3.3 | 908  | 4 | US-09-969-532-10     | Sequence 10, Appl  |
| 437 | 6 | 3.3 | 608 | 4 | US-08-957-310-21     | Sequence 21, Appl | 510 | 6 | 3.3 | 921  | 4 | US-09-543-681A-5734  | Sequence 5734, Ap  |
| 438 | 6 | 3.3 | 608 | 4 | US-10-011-366-21     | Sequence 21, Appl | 511 | 6 | 3.3 | 928  | 4 | US-09-134-000C-6590  | Sequence 6590, Ap  |
| 439 | 6 | 3.3 | 608 | 4 | US-09-084-517-21     | Sequence 21, Appl | 512 | 6 | 3.3 | 944  | 4 | US-09-341-505-1      | Sequence 1, Appl   |
| 440 | 6 | 3.3 | 609 | 1 | US-08-480-604A-30    | Sequence 30, Appl | 513 | 6 | 3.3 | 957  | 4 | US-09-252-991A-21567 | Sequence 21567, A  |
| 441 | 6 | 3.3 | 609 | 3 | US-08-915-136-30     | Sequence 30, Appl | 514 | 6 | 3.3 | 959  | 4 | US-09-232-991A-21747 | Sequence 21747, A  |
| 442 | 6 | 3.3 | 616 | 4 | US-09-084-517-30     | Sequence 30, Appl | 515 | 6 | 3.3 | 959  | 4 | US-09-328-352-7942   | Sequence 7942, Ap  |
| 443 | 6 | 3.3 | 631 | 4 | US-09-248-796A-20955 | Sequence 20955, A | 516 | 6 | 3.3 | 962  | 4 | US-09-328-352-7942   | Sequence 7942, Ap  |
| 444 | 6 | 3.3 | 631 | 4 | US-09-545-773-4      | Sequence 4, Appl  | 517 | 6 | 3.3 | 977  | 4 | US-09-134-000C-5653  | Sequence 5653, Ap  |
| 445 | 6 | 3.3 | 637 | 4 | US-09-248-796A-17906 | Sequence 17906, A | 518 | 6 | 3.3 | 980  | 4 | US-09-248-796A-19242 | Sequence 19242, A  |
| 446 | 6 | 3.3 | 638 | 4 | US-09-252-991A-25205 | Sequence 25205, A | 519 | 6 | 3.3 | 994  | 4 | US-09-252-991A-17781 | Sequence 17781, Ap |
| 447 | 6 | 3.3 | 644 | 4 | US-09-489-039A-12434 | Sequence 12434, A | 520 | 6 | 3.3 | 1028 | 4 | US-09-583-110-4290   | Sequence 4290, Ap  |
| 448 | 6 | 3.3 | 644 | 4 | US-09-248-796A-17325 | Sequence 17325, A | 521 | 6 | 3.3 | 1091 | 4 | US-09-538-092-469    | Sequence 469, App  |
| 449 | 6 | 3.3 | 646 | 4 | US-09-248-796A-16701 | Sequence 16701, A | 522 | 6 | 3.3 | 1116 | 4 | US-09-252-991A-24374 | Sequence 24374, A  |
| 450 | 6 | 3.3 | 648 | 3 | US-08-286-870A-4     | Sequence 4, Appl  | 523 | 6 | 3.3 | 1205 | 4 | US-09-252-991A-28876 | Sequence 28876, A  |
| 451 | 6 | 3.3 | 648 | 4 | US-09-328-352-4985   | Sequence 4985, Ap | 524 | 6 | 3.3 | 1253 | 3 | US-08-864-785-2      | Sequence 2, Appl   |
| 452 | 6 | 3.3 | 653 | 4 | US-09-252-991A-32608 | Sequence 32608, A | 525 | 6 | 3.3 | 1288 | 1 | US-07-727-814B-2     | Sequence 2, Appl   |
| 453 | 6 | 3.3 | 659 | 4 | US-09-328-352-6021   | Sequence 6021, Ap | 526 | 6 | 3.3 | 1288 | 1 | US-08-258-614-2      | Sequence 2, Appl   |
| 454 | 6 | 3.3 | 662 | 3 | US-09-405-728-3      | Sequence 3, Appl  | 527 | 6 | 3.3 | 1296 | 4 | US-08-857-636-60     | Sequence 60, Appl  |
| 455 | 6 | 3.3 | 672 | 1 | US-07-841-651-2      | Sequence 2, Appl  | 528 | 6 | 3.3 | 1398 | 1 | US-08-750-532-9      | Sequence 9, Appl   |
| 456 | 6 | 3.3 | 672 | 1 | US-07-841-651-3      | Sequence 3, Appl  | 529 | 6 | 3.3 | 1398 | 3 | US-08-894-818B-8     | Sequence 8, Appl   |
| 457 | 6 | 3.3 | 672 | 4 | US-09-253-991A-22808 | Sequence 22808, A | 530 | 6 | 3.3 | 1398 | 3 | US-09-445-472-6      | Sequence 6, Appl   |
| 458 | 6 | 3.3 | 672 | 4 | US-10-162-012-30     | Sequence 30, Appl | 531 | 6 | 3.3 | 1398 | 4 | US-10-090-624-6      | Sequence 81, Appl  |
| 459 | 6 | 3.3 | 673 | 4 | US-09-252-991A-23922 | Sequence 23922, A | 532 | 6 | 3.3 | 1422 | 4 | US-08-469-260A-81    | Sequence 81, Appl  |
| 460 | 6 | 3.3 | 681 | 4 | US-10-158-847-140    | Sequence 140, App | 533 | 6 | 3.3 | 1422 | 4 | US-08-467-344A-81    | Sequence 81, Appl  |
| 461 | 6 | 3.3 | 693 | 4 | US-09-252-991A-19167 | Sequence 19167, A | 534 | 6 | 3.3 | 1422 | 4 | US-08-424-550B-81    | Sequence 81, Appl  |
| 462 | 6 | 3.3 | 694 | 1 | US-08-171-382-2      | Sequence 2, Appl  | 535 | 6 | 3.3 | 1433 | 2 | US-08-365-486A-21    | Sequence 21, Appl  |
| 463 | 6 | 3.3 | 694 | 1 | US-08-309-420-2      | Sequence 2, Appl  | 536 | 6 | 3.3 | 1433 | 2 | US-09-123-708-4      | Sequence 4, Appl   |
| 464 | 6 | 3.3 | 694 | 1 | US-08-309-419-2      | Sequence 2, Appl  | 537 | 6 | 3.3 | 1433 | 3 | US-09-123-708-4      | Sequence 4, Appl   |
| 465 | 6 | 3.3 | 694 | 3 | US-09-294-531B-4     | Sequence 4, Appl  | 538 | 6 | 3.3 | 1433 | 3 | US-09-123-624-4      | Sequence 4, Appl   |

|     |   |     |      |   |                     |                    |     |   |     |    |   |                    |                    |
|-----|---|-----|------|---|---------------------|--------------------|-----|---|-----|----|---|--------------------|--------------------|
| 539 | 6 | 3.3 | 1433 | 3 | US-08-880-342-21    | Sequence 21, Appl  | 612 | 5 | 2.8 | 10 | 3 | US-07-978-674B-42  | Sequence 42, Appl  |
| 540 | 6 | 3.3 | 1434 | 2 | US-08-365-486A-19   | Sequence 19, Appl  | 613 | 5 | 2.8 | 10 | 3 | US-07-978-674B-43  | Sequence 43, Appl  |
| 541 | 6 | 3.3 | 1434 | 2 | US-08-540-406-10    | Sequence 10, Appl  | 614 | 5 | 2.8 | 10 | 3 | US-07-978-674B-44  | Sequence 44, Appl  |
| 542 | 6 | 3.3 | 1434 | 3 | US-08-656-055-10    | Sequence 10, Appl  | 615 | 5 | 2.8 | 10 | 3 | US-07-978-674B-45  | Sequence 45, Appl  |
| 543 | 6 | 3.3 | 1434 | 3 | US-08-954-668-10    | Sequence 10, Appl  | 616 | 5 | 2.8 | 10 | 3 | US-07-978-674B-46  | Sequence 46, Appl  |
| 544 | 6 | 3.3 | 1434 | 3 | US-08-880-342-19    | Sequence 19, Appl  | 617 | 5 | 2.8 | 10 | 3 | US-07-978-674B-47  | Sequence 47, Appl  |
| 545 | 6 | 3.3 | 1434 | 4 | US-08-918-658-10    | Sequence 10, Appl  | 618 | 5 | 2.8 | 10 | 4 | US-08-469-260A-278 | Sequence 278, Appl |
| 546 | 6 | 3.3 | 1434 | 4 | US-09-724-631-10    | Sequence 10, Appl  | 619 | 5 | 2.8 | 10 | 4 | US-08-488-446-278  | Sequence 278, Appl |
| 547 | 6 | 3.3 | 1434 | 4 | US-08-954-701A-10   | Sequence 10, Appl  | 620 | 5 | 2.8 | 10 | 4 | US-08-467-344A-278 | Sequence 278, Appl |
| 548 | 6 | 3.3 | 1434 | 4 | US-09-661-258-1     | Sequence 1, Appl   | 621 | 5 | 2.8 | 10 | 4 | US-08-424-550B-278 | Sequence 278, Appl |
| 549 | 6 | 3.3 | 1434 | 4 | US-09-538-092-1034  | Sequence 1034, Ap  | 622 | 5 | 2.8 | 11 | 2 | US-08-476-176B-53  | Sequence 53, Appl  |
| 550 | 6 | 3.3 | 1434 | 5 | PCT-US95-13233-10   | Sequence 10, Appl  | 623 | 5 | 2.8 | 11 | 2 | US-08-466-860-31   | Sequence 31, Appl  |
| 551 | 6 | 3.3 | 1447 | 2 | US-08-540-406-19    | Sequence 19, Appl  | 624 | 5 | 2.8 | 11 | 3 | US-08-127-721A-53  | Sequence 53, Appl  |
| 552 | 6 | 3.3 | 1447 | 3 | US-08-656-055-19    | Sequence 19, Appl  | 625 | 5 | 2.8 | 11 | 3 | US-08-485-246A-53  | Sequence 53, Appl  |
| 553 | 6 | 3.3 | 1447 | 3 | US-08-954-668-19    | Sequence 19, Appl  | 626 | 5 | 2.8 | 11 | 3 | US-08-472-040A-31  | Sequence 31, Appl  |
| 554 | 6 | 3.3 | 1447 | 3 | US-09-268-140-5     | Sequence 5, Appl   | 627 | 5 | 2.8 | 11 | 3 | US-08-276-776-31   | Sequence 31, Appl  |
| 555 | 6 | 3.3 | 1447 | 4 | US-08-918-658-19    | Sequence 19, Appl  | 628 | 5 | 2.8 | 11 | 3 | US-08-471-209-31   | Sequence 31, Appl  |
| 556 | 6 | 3.3 | 1447 | 4 | US-09-724-631-19    | Sequence 19, Appl  | 629 | 5 | 2.8 | 11 | 3 | US-08-977-378-7    | Sequence 7, Appl   |
| 557 | 6 | 3.3 | 1447 | 4 | US-08-954-701A-19   | Sequence 19, Appl  | 630 | 5 | 2.8 | 11 | 3 | US-08-977-378-15   | Sequence 15, Appl  |
| 558 | 6 | 3.3 | 1447 | 5 | PCT-US95-13233-19   | Sequence 19, Appl  | 631 | 5 | 2.8 | 11 | 3 | US-08-977-378-16   | Sequence 16, Appl  |
| 559 | 6 | 3.3 | 1554 | 2 | US-08-705-625-3     | Sequence 3, Appl   | 632 | 5 | 2.8 | 11 | 3 | US-08-977-378-23   | Sequence 23, Appl  |
| 560 | 6 | 3.3 | 1554 | 3 | US-09-010-998-6     | Sequence 6, Appl   | 633 | 5 | 2.8 | 11 | 3 | US-09-391-104-13   | Sequence 13, Appl  |
| 561 | 6 | 3.3 | 1554 | 3 | US-09-220-574-3     | Sequence 3, Appl   | 634 | 5 | 2.8 | 11 | 4 | US-08-055-006-10   | Sequence 10, Appl  |
| 562 | 6 | 3.3 | 1703 | 4 | US-09-824-574-3     | Sequence 3, Appl   | 635 | 5 | 2.8 | 11 | 4 | US-09-192-854-178  | Sequence 178, Appl |
| 563 | 6 | 3.3 | 2366 | 1 | US-08-480-604A-10   | Sequence 10, Appl  | 636 | 5 | 2.8 | 12 | 1 | US-07-616-910-47   | Sequence 47, Appl  |
| 564 | 6 | 3.3 | 2366 | 2 | US-08-405-496A-10   | Sequence 10, Appl  | 637 | 5 | 2.8 | 12 | 4 | US-09-446-787B-68  | Sequence 68, Appl  |
| 565 | 6 | 3.3 | 2366 | 2 | US-08-915-136-10    | Sequence 10, Appl  | 638 | 5 | 2.8 | 12 | 5 | PCT-US91-08497-47  | Sequence 47, Appl  |
| 566 | 6 | 3.3 | 2366 | 3 | US-08-957-310-10    | Sequence 10, Appl  | 639 | 5 | 2.8 | 13 | 2 | US-08-413-490-59   | Sequence 59, Appl  |
| 567 | 6 | 3.3 | 2366 | 4 | US-10-011-366-10    | Sequence 10, Appl  | 640 | 5 | 2.8 | 13 | 3 | US-09-146-269-59   | Sequence 59, Appl  |
| 568 | 6 | 3.3 | 2366 | 4 | US-09-084-517-10    | Sequence 10, Appl  | 641 | 5 | 2.8 | 13 | 3 | US-09-146-268-59   | Sequence 59, Appl  |
| 569 | 6 | 3.3 | 2813 | 3 | US-08-896-449A-2    | Sequence 2, Appl   | 642 | 5 | 2.8 | 13 | 4 | US-09-553-042-59   | Sequence 59, Appl  |
| 570 | 6 | 3.3 | 2813 | 3 | US-09-132-652-2     | Sequence 2, Appl   | 643 | 5 | 2.8 | 13 | 4 | US-09-704-090-1    | Sequence 1, Appl   |
| 571 | 6 | 3.3 | 2813 | 4 | US-09-886-900A-2    | Sequence 2, Appl   | 644 | 5 | 2.8 | 13 | 5 | PCT-US95-13037-59  | Sequence 59, Appl  |
| 572 | 6 | 3.3 | 2813 | 4 | US-09-662-478C-2    | Sequence 2, Appl   | 645 | 5 | 2.8 | 14 | 1 | US-08-205-719-5    | Sequence 5, Appl   |
| 573 | 6 | 3.3 | 2938 | 5 | PCT-US94-00198-3    | Sequence 3, Appl   | 646 | 5 | 2.8 | 14 | 2 | US-08-458-568A-10  | Sequence 10, Appl  |
| 574 | 6 | 3.3 | 3079 | 4 | US-09-091-609-2     | Sequence 2, Appl   | 647 | 5 | 2.8 | 14 | 3 | US-08-817-441-25   | Sequence 25, Appl  |
| 575 | 6 | 3.3 | 4630 | 4 | US-09-877-605-341   | Sequence 341, App  | 648 | 5 | 2.8 | 14 | 6 | 5245013-8          | Patent No. 5245013 |
| 576 | 6 | 3.3 | 5215 | 3 | US-09-105-537-2     | Sequence 2, Appl   | 649 | 5 | 2.8 | 15 | 3 | US-08-977-378-27   | Sequence 27, Appl  |
| 577 | 5 | 2.8 | 5215 | 6 | US-09-188-579-37    | Sequence 37, Appl  | 650 | 5 | 2.8 | 15 | 4 | US-09-386-959-61   | Sequence 61, Appl  |
| 578 | 5 | 2.8 | 6    | 3 | US-09-315-444-37    | Sequence 37, Appl  | 651 | 5 | 2.8 | 15 | 5 | PCT-US91-05177-3   | Sequence 3, Appl   |
| 579 | 5 | 2.8 | 6    | 3 | US-09-187-859-3997  | Sequence 3997, Ap  | 652 | 5 | 2.8 | 16 | 2 | US-08-889-291-16   | Sequence 16, Appl  |
| 580 | 5 | 2.8 | 6    | 4 | US-09-721-362-37    | Sequence 37, Appl  | 653 | 5 | 2.8 | 16 | 2 | US-09-100-414B-2   | Sequence 2, Appl   |
| 581 | 5 | 2.8 | 6    | 4 | US-09-752-165-75    | Sequence 75, Appl  | 654 | 5 | 2.8 | 16 | 3 | US-09-100-409A-61  | Sequence 61, Appl  |
| 582 | 5 | 2.8 | 6    | 4 | US-09-839-542B-3997 | Sequence 3997, Ap  | 655 | 5 | 2.8 | 16 | 3 | US-09-098-244-16   | Sequence 16, Appl  |
| 583 | 5 | 2.8 | 7    | 1 | US-08-462-661A-25   | Sequence 25, Appl  | 656 | 5 | 2.8 | 16 | 3 | US-08-602-998A-217 | Sequence 217, App  |
| 584 | 5 | 2.8 | 7    | 1 | US-08-462-661A-33   | Sequence 33, Appl  | 657 | 5 | 2.8 | 16 | 3 | US-08-602-998A-236 | Sequence 236, App  |
| 585 | 5 | 2.8 | 7    | 3 | US-09-264-709A-9    | Sequence 9, Appl   | 658 | 5 | 2.8 | 16 | 3 | US-09-303-323-2    | Sequence 2, Appl   |
| 586 | 5 | 2.8 | 7    | 4 | US-08-877-605-341   | Sequence 341, App  | 659 | 5 | 2.8 | 16 | 3 | US-09-375-314-16   | Sequence 16, Appl  |
| 587 | 5 | 2.8 | 7    | 6 | 5318899-46          | Patent No. 5318899 | 660 | 5 | 2.8 | 16 | 4 | US-09-500-124-217  | Sequence 217, App  |
| 588 | 5 | 2.8 | 7    | 6 | 5318899-52          | Sequence 130, App  | 661 | 5 | 2.8 | 16 | 4 | US-09-500-124-236  | Sequence 236, App  |
| 589 | 5 | 2.8 | 8    | 3 | US-07-861-458C-130  | Sequence 130, App  | 662 | 5 | 2.8 | 16 | 4 | US-09-767-395-16   | Sequence 16, Appl  |
| 590 | 5 | 2.8 | 8    | 3 | US-08-637-732A-34   | Sequence 34, Appl  | 663 | 5 | 2.8 | 16 | 4 | US-09-770-014-2    | Sequence 2, Appl   |
| 591 | 5 | 2.8 | 8    | 3 | US-08-637-732A-36   | Sequence 36, Appl  | 664 | 5 | 2.8 | 16 | 4 | US-09-701-588C-2   | Sequence 2, Appl   |
| 592 | 5 | 2.8 | 9    | 1 | US-08-487-890A-145  | Sequence 145, App  | 665 | 5 | 2.8 | 16 | 4 | US-09-747-802-31   | Sequence 31, Appl  |
| 593 | 5 | 2.8 | 9    | 2 | US-08-340-283-59    | Sequence 59, Appl  | 666 | 5 | 2.8 | 16 | 6 | 5200320-40         | Patent No. 5200320 |
| 594 | 5 | 2.8 | 9    | 2 | US-08-340-283-118   | Sequence 118, App  | 667 | 5 | 2.8 | 17 | 1 | US-08-395-381C-17  | Sequence 17, Appl  |
| 595 | 5 | 2.8 | 9    | 2 | US-08-340-283-132   | Sequence 132, App  | 668 | 5 | 2.8 | 17 | 3 | US-08-816-346-60   | Sequence 60, Appl  |
| 596 | 5 | 2.8 | 9    | 2 | US-08-478-435-145   | Sequence 145, App  | 669 | 5 | 2.8 | 17 | 3 | US-09-335-411-60   | Sequence 60, Appl  |
| 597 | 5 | 2.8 | 9    | 2 | US-08-337-483-145   | Sequence 145, App  | 670 | 5 | 2.8 | 17 | 3 | US-08-990-823-94   | Sequence 94, Appl  |
| 598 | 5 | 2.8 | 9    | 2 | US-08-478-373-145   | Sequence 145, App  | 671 | 5 | 2.8 | 17 | 3 | US-08-990-823-95   | Sequence 95, Appl  |
| 599 | 5 | 2.8 | 9    | 3 | US-08-474-671-145   | Sequence 145, App  | 672 | 5 | 2.8 | 17 | 4 | US-09-477-135A-94  | Sequence 94, Appl  |
| 600 | 5 | 2.8 | 9    | 3 | US-08-483-577A-145  | Sequence 145, App  | 673 | 5 | 2.8 | 17 | 4 | US-09-477-135A-95  | Sequence 95, Appl  |
| 601 | 5 | 2.8 | 9    | 3 | US-08-997-438-145   | Sequence 145, App  | 674 | 5 | 2.8 | 17 | 4 | US-09-326-447-17   | Sequence 17, Appl  |
| 602 | 5 | 2.8 | 9    | 3 | US-08-637-654-145   | Sequence 145, App  | 675 | 5 | 2.8 | 17 | 4 | US-09-493-795B-183 | Sequence 183, App  |
| 603 | 5 | 2.8 | 9    | 3 | US-08-649-518-145   | Sequence 145, App  | 676 | 5 | 2.8 | 19 | 1 | US-08-664-596B-8   | Sequence 8, Appl   |
| 604 | 5 | 2.8 | 9    | 4 | US-09-647-372B-5    | Sequence 5, Appl   | 677 | 5 | 2.8 | 19 | 2 | US-08-888-497-4    | Sequence 4, Appl   |
| 605 | 5 | 2.8 | 10   | 2 | US-08-724-548-42    | Sequence 42, Appl  | 678 | 5 | 2.8 | 19 | 3 | US-09-169-015-52   | Sequence 52, Appl  |
| 606 | 5 | 2.8 | 10   | 2 | US-08-724-548-43    | Sequence 43, Appl  | 679 | 5 | 2.8 | 19 | 3 | US-09-362-230-4    | Sequence 4, Appl   |
| 607 | 5 | 2.8 | 10   | 2 | US-08-724-548-44    | Sequence 44, Appl  | 680 | 5 | 2.8 | 19 | 4 | US-09-626-581D-57  | Sequence 57, Appl  |
| 608 | 5 | 2.8 | 10   | 2 | US-08-724-548-45    | Sequence 45, Appl  | 681 | 5 | 2.8 | 19 | 4 | US-09-415-765B-57  | Sequence 57, Appl  |
| 609 | 5 | 2.8 | 10   | 2 | US-08-724-548-46    | Sequence 46, Appl  | 682 | 5 | 2.8 | 19 | 4 | US-09-626-580C-57  | Sequence 57, Appl  |
| 610 | 5 | 2.8 | 10   | 2 | US-08-724-548-47    | Sequence 47, Appl  | 683 | 5 | 2.8 | 19 | 4 | US-09-749-959-49   | Sequence 49, Appl  |
| 611 | 5 | 2.8 | 10   | 3 | US-08-297-395-5     | Sequence 5, Appl   | 684 | 5 | 2.8 | 19 | 4 | US-09-178-093B-21  | Sequence 21, Appl  |

|     |   |     |    |   |                     |                    |     |   |     |    |   |                     |                    |
|-----|---|-----|----|---|---------------------|--------------------|-----|---|-----|----|---|---------------------|--------------------|
| 685 | 5 | 2.8 | 19 | 5 | PCT-US94-07926-4    | Sequence 4, Appli  | 758 | 5 | 2.8 | 28 | 3 | US-09-058-459-11    | Sequence 11, Appli |
| 686 | 5 | 2.8 | 20 | 2 | US-08-888-497-9     | Sequence 9, Appli  | 759 | 5 | 2.8 | 28 | 3 | US-09-058-459-12    | Sequence 12, Appli |
| 687 | 5 | 2.8 | 20 | 2 | US-08-888-497-11    | Sequence 11, Appli | 760 | 5 | 2.8 | 28 | 3 | US-09-058-459-13    | Sequence 13, Appli |
| 688 | 5 | 2.8 | 20 | 3 | US-09-166-028-5     | Sequence 5, Appli  | 761 | 5 | 2.8 | 28 | 3 | US-09-058-459-14    | Sequence 14, Appli |
| 689 | 5 | 2.8 | 20 | 3 | US-09-362-230-9     | Sequence 11, Appli | 762 | 5 | 2.8 | 28 | 3 | US-09-058-459-15    | Sequence 15, Appli |
| 690 | 5 | 2.8 | 20 | 3 | US-09-362-230-11    | Sequence 11, Appli | 763 | 5 | 2.8 | 28 | 3 | US-09-058-459-16    | Sequence 16, Appli |
| 691 | 5 | 2.8 | 20 | 4 | US-08-465-260A-291  | Sequence 291, App  | 764 | 5 | 2.8 | 28 | 3 | US-09-058-459-19    | Sequence 21, Appli |
| 692 | 5 | 2.8 | 20 | 4 | US-08-488-446-291   | Sequence 291, App  | 765 | 5 | 2.8 | 28 | 3 | US-09-058-459-21    | Sequence 21, Appli |
| 693 | 5 | 2.8 | 20 | 4 | US-08-467-344A-291  | Sequence 291, App  | 766 | 5 | 2.8 | 28 | 3 | US-09-093-539-95    | Sequence 37, Appli |
| 694 | 5 | 2.8 | 20 | 4 | US-08-424-550B-291  | Sequence 291, App  | 767 | 5 | 2.8 | 28 | 3 | US-09-303-323-37    | Sequence 2, Appli  |
| 695 | 5 | 2.8 | 20 | 5 | PCT-US94-07926-9    | Sequence 9, Appli  | 768 | 5 | 2.8 | 28 | 3 | US-09-127-926-2     | Sequence 2, Appli  |
| 696 | 5 | 2.8 | 20 | 5 | PCT-US94-07926-11   | Sequence 11, Appli | 769 | 5 | 2.8 | 28 | 3 | US-09-127-926-3     | Sequence 3, Appli  |
| 697 | 5 | 2.8 | 21 | 2 | US-08-679-635A-8    | Sequence 8, Appli  | 770 | 5 | 2.8 | 28 | 3 | US-09-127-926-11    | Sequence 11, Appli |
| 698 | 5 | 2.8 | 21 | 3 | US-09-419-163-8     | Sequence 8, Appli  | 771 | 5 | 2.8 | 28 | 3 | US-09-127-926-12    | Sequence 12, Appli |
| 699 | 5 | 2.8 | 21 | 3 | US-09-409-604-17    | Sequence 17, Appli | 772 | 5 | 2.8 | 28 | 3 | US-09-127-926-14    | Sequence 14, Appli |
| 700 | 5 | 2.8 | 21 | 6 | 5245013-9           | Patent No. 5245013 | 773 | 5 | 2.8 | 28 | 3 | US-09-127-926-15    | Sequence 15, Appli |
| 701 | 5 | 2.8 | 22 | 5 | PCT-US92-08094-5    | Sequence 5, Appli  | 774 | 5 | 2.8 | 28 | 3 | US-09-127-926-16    | Sequence 16, Appli |
| 702 | 5 | 2.8 | 23 | 4 | US-08-981-527A-19   | Sequence 19, Appli | 775 | 5 | 2.8 | 28 | 3 | US-03-217-352-89    | Sequence 89, Appli |
| 703 | 5 | 2.8 | 23 | 4 | US-10-147-936B-19   | Sequence 19, Appli | 776 | 5 | 2.8 | 28 | 3 | US-09-432-879-8     | Sequence 8, Appli  |
| 704 | 5 | 2.8 | 24 | 1 | US-07-670-611-11    | Sequence 11, Appli | 777 | 5 | 2.8 | 28 | 4 | US-09-790-230-95    | Sequence 95, Appli |
| 705 | 5 | 2.8 | 24 | 1 | US-08-220-674-11    | Sequence 11, Appli | 778 | 5 | 2.8 | 28 | 4 | US-09-770-014-37    | Sequence 37, Appli |
| 706 | 5 | 2.8 | 24 | 1 | US-08-445-186-11    | Sequence 11, Appli | 779 | 5 | 2.8 | 28 | 4 | US-09-677-664B-33   | Sequence 33, Appli |
| 707 | 5 | 2.8 | 24 | 1 | US-08-446-549-11    | Sequence 11, Appli | 780 | 5 | 2.8 | 28 | 4 | US-09-689-097-96    | Sequence 96, Appli |
| 708 | 5 | 2.8 | 24 | 1 | US-08-461-597-4     | Sequence 4, Appli  | 781 | 5 | 2.8 | 28 | 4 | US-09-689-097-96    | Sequence 96, Appli |
| 709 | 5 | 2.8 | 24 | 2 | US-08-446-550-11    | Sequence 11, Appli | 782 | 5 | 2.8 | 28 | 4 | US-09-270-767-60554 | Sequence 60554, A  |
| 710 | 5 | 2.8 | 24 | 2 | US-08-535-298-4     | Sequence 4, Appli  | 783 | 5 | 2.8 | 28 | 4 | US-09-714-357-2     | Sequence 2, Appli  |
| 711 | 5 | 2.8 | 24 | 4 | US-09-428-082B-1050 | Sequence 1050, Ap  | 784 | 5 | 2.8 | 28 | 4 | US-09-714-357-3     | Sequence 3, Appli  |
| 712 | 5 | 2.8 | 24 | 4 | US-09-721-108-226   | Sequence 226, App  | 785 | 5 | 2.8 | 28 | 4 | US-09-714-357-7     | Sequence 7, Appli  |
| 713 | 5 | 2.8 | 24 | 4 | US-09-721-108-227   | Sequence 227, App  | 786 | 5 | 2.8 | 28 | 4 | US-09-714-357-11    | Sequence 11, Appli |
| 714 | 5 | 2.8 | 24 | 4 | US-09-270-767-39936 | Sequence 227, App  | 787 | 5 | 2.8 | 28 | 4 | US-09-714-357-12    | Sequence 12, Appli |
| 715 | 5 | 2.8 | 24 | 4 | US-09-270-767-55153 | Sequence 39936, A  | 788 | 5 | 2.8 | 28 | 4 | US-09-714-357-13    | Sequence 13, Appli |
| 716 | 5 | 2.8 | 24 | 5 | PCT-US94-05569A-4   | Sequence 4, Appli  | 789 | 5 | 2.8 | 28 | 4 | US-09-714-357-14    | Sequence 14, Appli |
| 717 | 5 | 2.8 | 25 | 1 | US-08-345-527-5     | Sequence 4, Appli  | 790 | 5 | 2.8 | 28 | 4 | US-09-714-357-15    | Sequence 15, Appli |
| 718 | 5 | 2.8 | 25 | 1 | US-08-345-527-5     | Sequence 5, Appli  | 791 | 5 | 2.8 | 28 | 4 | US-09-714-357-16    | Sequence 16, Appli |
| 719 | 5 | 2.8 | 25 | 2 | US-08-448-600-8     | Sequence 8, Appli  | 792 | 5 | 2.8 | 28 | 4 | US-09-714-357-19    | Sequence 19, Appli |
| 720 | 5 | 2.8 | 25 | 3 | US-08-887-769-5     | Sequence 8, Appli  | 793 | 5 | 2.8 | 28 | 4 | US-09-714-357-21    | Sequence 21, Appli |
| 721 | 5 | 2.8 | 25 | 3 | US-08-937-610-15    | Sequence 15, Appli | 794 | 5 | 2.8 | 28 | 4 | US-09-701-588C-37   | Sequence 37, Appli |
| 722 | 5 | 2.8 | 25 | 4 | US-09-270-767-41088 | Sequence 41088, A  | 795 | 5 | 2.8 | 28 | 4 | US-10-057-552-2     | Sequence 2, Appli  |
| 723 | 5 | 2.8 | 25 | 4 | US-09-270-767-56304 | Sequence 56304, A  | 796 | 5 | 2.8 | 28 | 4 | US-10-057-552-3     | Sequence 3, Appli  |
| 724 | 5 | 2.8 | 26 | 6 | 5245013-19          | Patent No. 5245013 | 797 | 5 | 2.8 | 28 | 4 | US-10-057-552-7     | Sequence 7, Appli  |
| 725 | 5 | 2.8 | 26 | 1 | US-08-446-692-80    | Sequence 80, Appli | 798 | 5 | 2.8 | 28 | 4 | US-10-057-552-11    | Sequence 11, Appli |
| 726 | 5 | 2.8 | 26 | 2 | US-08-488-351A-80   | Sequence 80, Appli | 799 | 5 | 2.8 | 28 | 4 | US-10-057-552-12    | Sequence 12, Appli |
| 727 | 5 | 2.8 | 26 | 4 | US-09-461-325-434   | Sequence 434, App  | 800 | 5 | 2.8 | 28 | 4 | US-10-057-552-13    | Sequence 13, Appli |
| 728 | 5 | 2.8 | 26 | 4 | US-10-012-542-434   | Sequence 434, App  | 801 | 5 | 2.8 | 28 | 4 | US-10-057-552-14    | Sequence 14, Appli |
| 729 | 5 | 2.8 | 26 | 4 | US-09-270-767-40133 | Sequence 40133, A  | 802 | 5 | 2.8 | 28 | 4 | US-10-057-552-15    | Sequence 15, Appli |
| 730 | 5 | 2.8 | 26 | 4 | US-09-270-767-55349 | Sequence 55349, A  | 803 | 5 | 2.8 | 28 | 4 | US-10-057-552-16    | Sequence 16, Appli |
| 731 | 5 | 2.8 | 26 | 4 | US-09-270-767-59088 | Sequence 59088, A  | 804 | 5 | 2.8 | 28 | 4 | US-10-057-552-19    | Sequence 19, Appli |
| 732 | 5 | 2.8 | 26 | 4 | US-10-115-143-434   | Sequence 434, App  | 805 | 5 | 2.8 | 28 | 4 | US-10-057-552-21    | Sequence 21, Appli |
| 733 | 5 | 2.8 | 27 | 3 | US-09-127-926-13    | Sequence 13, Appli | 806 | 5 | 2.8 | 28 | 5 | PCT-US94-02465-95   | Sequence 95, Appli |
| 734 | 5 | 2.8 | 27 | 3 | US-09-507-819-24    | Sequence 24, Appli | 807 | 5 | 2.8 | 28 | 5 | PCT-US95-00498-95   | Sequence 95, Appli |
| 735 | 5 | 2.8 | 27 | 4 | US-09-641-576-24    | Sequence 24, Appli | 808 | 5 | 2.8 | 28 | 5 | PCT-US95-00656-95   | Sequence 95, Appli |
| 736 | 5 | 2.8 | 27 | 4 | US-09-270-767-39761 | Sequence 39761, A  | 809 | 5 | 2.8 | 28 | 6 | PCT-US95-09262-33   | Sequence 33, Appli |
| 737 | 5 | 2.8 | 27 | 4 | US-09-270-767-54978 | Sequence 54978, A  | 810 | 5 | 2.8 | 28 | 6 | 5171680-9           | Patent No. 5171680 |
| 738 | 5 | 2.8 | 27 | 4 | US-09-270-767-60216 | Sequence 60216, A  | 811 | 5 | 2.8 | 29 | 1 | US-07-741-940-9     | Sequence 9, Appli  |
| 739 | 5 | 2.8 | 28 | 1 | US-08-311-611A-95   | Sequence 95, Appli | 812 | 5 | 2.8 | 29 | 1 | US-08-289-548A-9    | Sequence 9, Appli  |
| 740 | 5 | 2.8 | 28 | 1 | US-08-372-783-95    | Sequence 95, Appli | 813 | 5 | 2.8 | 29 | 1 | US-08-452-654-9     | Sequence 9, Appli  |
| 741 | 5 | 2.8 | 28 | 1 | US-08-372-105-95    | Sequence 95, Appli | 814 | 5 | 2.8 | 29 | 1 | US-08-452-655B-9    | Sequence 9, Appli  |
| 742 | 5 | 2.8 | 28 | 1 | US-08-306-473A-95   | Sequence 95, Appli | 815 | 5 | 2.8 | 29 | 3 | US-08-450-582-9     | Sequence 9, Appli  |
| 743 | 5 | 2.8 | 28 | 1 | US-08-209-762-95    | Sequence 95, Appli | 816 | 5 | 2.8 | 29 | 3 | US-09-060-756-732   | Sequence 732, App  |
| 744 | 5 | 2.8 | 28 | 1 | US-08-473-344-95    | Sequence 95, Appli | 817 | 5 | 2.8 | 29 | 4 | US-08-449-731-9     | Sequence 9, Appli  |
| 745 | 5 | 2.8 | 28 | 1 | US-08-621-803-89    | Sequence 89, Appli | 818 | 5 | 2.8 | 29 | 4 | US-09-670-314-732   | Sequence 732, App  |
| 746 | 5 | 2.8 | 28 | 2 | US-08-485-445A-95   | Sequence 89, Appli | 819 | 5 | 2.8 | 29 | 4 | US-09-270-767-38456 | Sequence 38456, A  |
| 747 | 5 | 2.8 | 28 | 2 | US-08-621-259A-33   | Sequence 33, Appli | 820 | 5 | 2.8 | 29 | 4 | US-09-270-767-53673 | Sequence 53673, A  |
| 748 | 5 | 2.8 | 28 | 2 | US-08-635-007-8     | Sequence 8, Appli  | 821 | 5 | 2.8 | 29 | 4 | US-09-270-767-56905 | Sequence 56905, A  |
| 749 | 5 | 2.8 | 28 | 3 | US-09-110-953-8     | Sequence 8, Appli  | 822 | 5 | 2.8 | 30 | 2 | US-08-635-007-9     | Sequence 9, Appli  |
| 750 | 5 | 2.8 | 28 | 3 | US-09-100-414B-37   | Sequence 37, Appli | 823 | 5 | 2.8 | 30 | 2 | US-08-935-886-2     | Sequence 2, Appli  |
| 751 | 5 | 2.8 | 28 | 3 | US-09-235-246-4     | Sequence 4, Appli  | 824 | 5 | 2.8 | 30 | 3 | US-09-110-953-9     | Sequence 9, Appli  |
| 752 | 5 | 2.8 | 28 | 3 | US-09-119-263-95    | Sequence 95, Appli | 825 | 5 | 2.8 | 30 | 3 | US-09-432-879-9     | Sequence 9, Appli  |
| 753 | 5 | 2.8 | 28 | 3 | US-08-657-162-95    | Sequence 95, Appli | 826 | 5 | 2.8 | 31 | 3 | US-08-602-999A-43   | Sequence 43, Appli |
| 754 | 5 | 2.8 | 28 | 3 | US-09-224-480-95    | Sequence 95, Appli | 827 | 5 | 2.8 | 31 | 3 | US-08-278-865-43    | Sequence 43, Appli |
| 755 | 5 | 2.8 | 28 | 3 | US-09-058-459-2     | Sequence 2, Appli  | 828 | 5 | 2.8 | 31 | 4 | US-09-500-124-43    | Sequence 43, Appli |
| 756 | 5 | 2.8 | 28 | 3 | US-09-058-459-3     | Sequence 3, Appli  | 829 | 5 | 2.8 | 31 | 4 | US-09-938-315-43    | Sequence 43, Appli |
| 757 | 5 | 2.8 | 28 | 3 | US-09-058-459-7     | Sequence 7, Appli  | 830 | 5 | 2.8 | 31 | 4 | US-09-270-767-36287 | Sequence 36287, A  |
| 758 | 5 | 2.8 | 28 | 3 | US-09-058-459-11    | Sequence 11, Appli | 830 | 5 | 2.8 | 31 | 4 | US-09-270-767-51504 | Sequence 51504, A  |

831 5 2.8 31 4 US-09-270-767-59771 Sequence 59771, A  
832 5 2.8 31 6 5245013-17 Patent No. 5245013  
833 5 2.8 32 4 US-09-270-767-39951 Sequence 39951, A  
834 5 2.8 32 4 US-09-270-767-55168 Sequence 55168, A  
835 5 2.8 32 1 US-07-973-321B-10 Sequence 10, Appl  
836 5 2.8 33 1 US-08-090-527A-10 Sequence 10, Appl  
837 5 2.8 33 2 US-08-415-788-28 Sequence 28, Appl  
838 5 2.8 33 4 US-08-997-685A-51 Sequence 51, Appl  
839 5 2.8 33 4 US-09-086-436-58 Sequence 58, Appl  
840 5 2.8 33 4 US-08-624-735B-15 Sequence 15, Appl  
841 5 2.8 34 4 US-08-624-735B-16 Sequence 16, Appl  
842 5 2.8 34 2 US-08-415-788-11 Sequence 11, Appl  
843 5 2.8 34 2 US-08-415-788-14 Sequence 14, Appl  
844 5 2.8 34 2 US-08-415-788-18 Sequence 18, Appl  
845 5 2.8 34 2 US-08-415-788-24 Sequence 24, Appl  
846 5 2.8 34 3 US-08-967-508-1 Sequence 1, Appl  
847 5 2.8 34 3 US-08-967-508-1 Sequence 1, Appl  
848 5 2.8 34 4 US-09-270-767-34855 Sequence 34855, A  
849 5 2.8 34 4 US-09-270-767-50072 Sequence 50072, A  
850 5 2.8 34 5 PCT-US94-02552-1 Sequence 1, Appl  
851 5 2.8 37 1 US-08-164-151-14 Sequence 14, Appl  
852 5 2.8 37 2 US-08-448-481-5 Sequence 5, Appl  
853 5 2.8 37 4 US-09-461-325-265 Sequence 265, App  
854 5 2.8 37 4 US-09-673-395A-186 Sequence 186, App  
855 5 2.8 37 4 US-10-012-542-265 Sequence 265, App  
856 5 2.8 37 4 US-08-817-441-27 Sequence 27, Appl  
857 5 2.8 38 3 US-08-817-441-93 Sequence 93, Appl  
858 5 2.8 38 4 US-09-285-912A-68 Sequence 68, Appl  
859 5 2.8 39 4 US-09-270-767-40898 Sequence 40898, A  
860 5 2.8 39 4 US-09-270-767-55914 Sequence 55914, A  
861 5 2.8 39 4 US-09-493-795B-308 Sequence 308, App  
862 5 2.8 39 6 5245013-1 Patent No. 5245013  
863 5 2.8 40 3 US-08-685-558A-1 Sequence 1, Appl  
864 5 2.8 40 4 US-09-765-449-1 Sequence 1, Appl  
865 5 2.8 41 4 US-09-673-395A-467 Sequence 467, App  
866 5 2.8 41 4 US-08-377-687-39 Sequence 39, Appl  
867 5 2.8 42 1 US-08-777-192-39 Sequence 39, Appl  
868 5 2.8 42 2 US-08-488-161-54 Sequence 54, Appl  
869 5 2.8 42 4 US-09-273-685-54 Sequence 54, Appl  
870 5 2.8 42 3 US-08-971-982-39 Sequence 39, Appl  
871 5 2.8 42 4 US-09-270-767-34493 Sequence 34493, A  
872 5 2.8 42 4 US-09-270-767-49710 Sequence 49710, A  
873 5 2.8 42 4 US-09-270-767-58998 Sequence 58998, A  
874 5 2.8 42 4 PCT-US95-11934-54 Sequence 54, Appl  
875 5 2.8 43 4 US-09-270-767-38449 Sequence 38449, A  
876 5 2.8 43 4 US-09-270-767-53666 Sequence 53666, A  
877 5 2.8 44 4 US-09-823-266-17 Sequence 17, Appl  
878 5 2.8 44 4 US-09-443-780C-19 Sequence 19, Appl  
879 5 2.8 44 6 5274075-12 Patent No. 5274075  
880 5 2.8 45 6 5274075-12 Patent No. 5274075  
881 5 2.8 45 2 US-08-691-814B-40 Sequence 40, Appl  
882 5 2.8 46 3 US-08-905-223-479 Sequence 479, App  
883 5 2.8 46 3 US-09-227-357-582 Sequence 582, App  
884 5 2.8 46 3 US-09-500-063-4 Sequence 4, Appl  
885 5 2.8 46 4 US-09-270-767-56966 Sequence 56966, A  
886 5 2.8 46 4 US-08-691-814B-39 Sequence 39, Appl  
887 5 2.8 47 2 US-09-515-965A-1943 Sequence 1943, App  
888 5 2.8 47 4 US-09-350-641C-1706 Sequence 1706, App  
889 5 2.8 47 4 US-08-743-200-8 Sequence 8, Appl  
890 5 2.8 48 2 US-08-460-890A-49 Sequence 49, Appl  
891 5 2.8 48 3 US-08-167-641C-49 Sequence 49, Appl  
892 5 2.8 48 3 US-08-460-890A-49 Sequence 49, Appl  
893 5 2.8 48 3 US-08-462-040-49 Sequence 49, Appl  
894 5 2.8 48 4 US-09-270-767-38458 Sequence 38458, A  
895 5 2.8 48 4 US-09-270-767-53675 Sequence 53675, A  
896 5 2.8 48 4 US-09-270-767-59066 Sequence 59066, A  
897 5 2.8 49 2 US-08-167-641C-48 Sequence 48, Appl  
898 5 2.8 49 3 US-08-460-890A-48 Sequence 48, Appl  
899 5 2.8 49 3 US-08-460-890A-48 Sequence 48, Appl  
900 5 2.8 49 3 US-08-460-890A-48 Sequence 48, Appl  
901 5 2.8 49 3 US-08-460-890A-48 Sequence 48, Appl  
902 5 2.8 49 3 US-08-460-890A-48 Sequence 48, Appl  
903 5 2.8 50 4 US-09-621-976-7049 Sequence 7049, App

50 4 US-08-980-326-68 Sequence 68, Appl  
51 4 US-09-270-767-35778 Sequence 35778, A  
52 4 US-09-270-767-50995 Sequence 50995, A  
53 4 US-09-270-767-61596 Sequence 61596, A  
54 4 US-09-270-767-61596 Sequence 61596, A  
55 4 US-09-270-767-61844 Patent No. 5274075  
56 6 5274075-6 Patent No. 5274075  
57 2 US-08-821-619-3 Sequence 3, Appl  
58 3 US-08-630-916A-115 Sequence 115, App  
59 3 US-09-209-525-55 Sequence 55, Appl  
60 3 US-08-309-420-7 Sequence 7, Appl  
61 3 US-08-309-420-7 Sequence 7, Appl  
62 2 US-08-470-419-4 Sequence 4, Appl  
63 2 US-08-761-828-4 Sequence 4, Appl  
64 2 US-08-290-105-4 Sequence 4, Appl  
65 2 US-08-776-949-4 Sequence 4, Appl  
66 3 US-08-482-810-4 Sequence 4, Appl  
67 3 US-09-027-955-4 Sequence 4, Appl  
68 3 US-09-636-808-4 Sequence 4, Appl  
69 3 US-09-258-128-4 Sequence 4, Appl  
70 4 US-09-635-754-4 Sequence 4, Appl  
71 4 US-08-680-525-4 Sequence 4, Appl  
72 4 US-09-636-223-4 Sequence 4, Appl  
73 5 PCT-US95-11856-7 PCT-US95-11856-7  
74 5 PCT-US95-11878-7 PCT-US95-11878-7  
75 2 US-08-746-283-9 Sequence 7, Appl  
76 4 US-08-746-257A-7 Sequence 7, Appl  
77 4 US-09-621-976-4897 Sequence 4897, App  
78 5 US-08-366-953A-36 Sequence 36, Appl  
79 5 US-08-366-953A-37 Sequence 37, Appl  
80 5 US-08-366-953A-38 Sequence 38, Appl  
81 5 US-08-366-953A-39 Sequence 39, Appl  
82 5 US-08-366-953A-40 Sequence 40, Appl  
83 5 US-08-366-953A-41 Sequence 41, Appl  
84 5 US-08-366-953A-42 Sequence 42, Appl  
85 5 US-08-366-953A-43 Sequence 43, Appl  
86 5 US-08-366-953A-44 Sequence 44, Appl  
87 5 US-09-513-999C-6973 Sequence 6973, App  
88 4 US-09-621-976-7643 Sequence 7643, App  
89 5 US-09-270-767-35043 Sequence 35043, A  
90 5 US-09-270-767-50260 Sequence 50260, A  
91 3 US-08-984-151-2 Sequence 2, Appl  
92 4 US-09-621-976-6016 Sequence 6016, App  
93 4 US-09-513-999C-7163 Sequence 7163, App  
94 5 US-08-676-125A-26 Sequence 26, Appl  
95 2 US-09-136-012A-26 Sequence 26, Appl  
96 4 US-09-441-340-15 Sequence 15, Appl  
97 4 US-09-513-999C-4594 Sequence 4594, App  
98 4 US-09-186-002-8 Sequence 8, Appl  
99 4 US-09-795-926-14 Sequence 14, Appl  
100 4 US-09-621-976-4744 Sequence 4744, App

Sequence 4744, App  
Sequence 6789, App  
Sequence 58655, A  
Sequence 4390, App  
Sequence 28, Appl  
Sequence 29, Appl  
Sequence 30, Appl  
Sequence 31, Appl  
Sequence 32, Appl  
Sequence 33, Appl  
Sequence 34, Appl  
Sequence 35, Appl  
Sequence 36, Appl  
Sequence 37, Appl  
Sequence 37A-37  
Sequence 5323, App  
Sequence 60124, A  
Sequence 23346, A  
Sequence 25581, A  
Sequence 27958, A  
Sequence 8137, App  
Sequence 4027, App  
Sequence 21014, A  
Sequence 22949, A  
Sequence 4208, App

```
977 5 2.8 61 4 US-09-513-999C-6266 Sequence 6266, Ap
978 5 2.8 62 4 US-09-205-258-470 Sequence 470, App
979 5 2.8 62 4 US-09-107-532A-7256 Sequence 7256, Ap
980 5 2.8 62 4 US-09-543-681A-7584 Sequence 7584, Ap
981 5 2.8 62 4 US-09-134-000C-3955 Sequence 3955, Ap
982 5 2.8 63 1 US-08-446-692-105 Sequence 105, App
983 5 2.8 63 2 US-08-488-351A-105 Sequence 105, App
984 5 2.8 63 4 US-09-540-236-2372 Sequence 2372, Ap
985 5 2.8 63 4 US-09-583-110-3574 Sequence 3574, Ap
986 5 2.8 63 4 US-09-248-796A-27026 Sequence 27026, A
987 5 2.8 64 3 US-08-753-463-4 Sequence 4, Appli
988 5 2.8 64 3 US-09-328-153-4 Sequence 4, Appli
989 5 2.8 64 3 US-09-227-357-243 Sequence 243, App
990 5 2.8 64 3 US-09-227-357-640 Sequence 640, App
991 5 2.8 64 4 US-09-489-039A-8132 Sequence 8132, Ap
992 5 2.8 64 4 US-09-621-976-5513 Sequence 5513, Ap
993 5 2.8 64 4 US-09-583-110-3389 Sequence 3389, Ap
994 5 2.8 64 4 US-09-583-110-3932 Sequence 3932, Ap
995 5 2.8 64 4 US-09-583-110-4338 Sequence 4338, Ap
996 5 2.8 64 4 US-09-270-767-36770 Sequence 36770, A
997 5 2.8 64 4 US-09-270-767-51987 Sequence 51987, A
998 5 2.8 64 4 US-09-270-767-60813 Sequence 60813, A
999 5 2.8 64 4 US-09-248-796A-22749 Sequence 22749, A
1000 5 2.8 64 4 US-09-248-796A-23728 Sequence 23728, A

ALIGNMENTS

RESULT 1
US-09-328-352-6050
; Sequence 6050, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6050
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6050

Query Match 8.9%; Score 16; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 8.3e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 PLPLGGGGGARIACGVI 178
Db 197 PLPLGGGGGARIACGVI 212
|||||

RESULT 2
US-07-641-346B-2
; Sequence 2, Application US/07641346B
; Patent No. 5188936
; GENERAL INFORMATION:
; APPLICANT: LOUISA B. TABATABAI
; APPLICANT: JOHN E. MAYFIELD
; TITLE OF INVENTION: BRUCELLA ABORTUS DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS P. RIBANDO
; ADDRESSEE: USDA-ARS-OCI
; ADDRESSEE: NATIONAL CENTER FOR AGRICULTURAL
; ADDRESSEE: UTILIZATION RESEARCH
; STREET: 1815 NORTH UNIVERSITY STREET
; CITY: PEORIA
```

```
; STATE: IL
; COUNTRY: U.S.A.
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE - 5.25 INCH, 360 Kb STORAGE
; COMPUTER: IBM XT COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: MULTIMATE ADVANTAGE II
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/641,346B
; FILING DATE: 19910116
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: N/A
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: RIBANDO, CURTIS P.
; REGISTRATION NUMBER: 27,976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: FTS 360-4513, COM 309/685-4011, x513
; TELEFAX: COM 309/685-4128
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-641-346B-2

Query Match 5.0%; Score 9; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDKP 163
Db 1 GGDNYSDKP 9
|||||

RESULT 3
US-09-489-039A-9511
; Sequence 9511, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9511
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9511

Query Match 5.0%; Score 9; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 TGKHEGPG 113
Db 69 TGKHEGPG 77
|||||

RESULT 4
US-07-641-346B-1
; Sequence 1, Application US/07641346B
; Patent No. 5188936
; GENERAL INFORMATION:
```

APPLICANT: LOUISA B. TABATABAI  
APPLICANT: JOHN E. MAYFIELD  
TITLE OF INVENTION: BRUCELLA ABORTUS DIAGNOSTIC  
TITLE OF INVENTION: REAGENTS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS P. RIBANDO  
ADDRESSEE: USDA-ARS-OCI  
ADDRESSEE: NATIONAL CENTER FOR AGRICULTURAL  
ADDRESSEE: UTILIZATION RESEARCH  
STREET: 1815 NORTH UNIVERSITY STREET  
CITY: PEORIA  
STATE: IL  
COUNTRY: U.S.A.  
ZIP: 61604  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE - 5.25 INCH, 360 Kb STORAGE  
COMPUTER: IBM XT COMPATIBLE  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: MULTIMATE ADVANTAGE II  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/641,346B  
FILING DATE: 19910116  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: N/A  
APPLICATION NUMBER: N/A  
FILING DATE: N/A  
ATTORNEY/AGENT INFORMATION:  
NAME: RIBANDO, CURTIS P.  
REGISTRATION NUMBER: 27,976  
TELEPHONE: FTS 360-4513, COM 309/685-4011, x513  
TELEFAX: COM 309/685-4128  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 154  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-641-346B-1

Query Match 5.0%; Score 9; DB 1; Length 154;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDYSDKP 163  
DB 130 GGDYSDKP 138  
RESULT 5  
US-08-676-967-1  
Sequence 1, Application US/08676967  
Patent No. 5747317  
GENERAL INFORMATION:  
APPLICANT: COLLINS, KATHLEEN  
TITLE OF INVENTION: Human Telomerase  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Science & Technology Law Group  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/676,967

FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UCB96-055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415)343-4342  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 759 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-676-967-1  
Query Match 4.4%; Score 8; DB 1; Length 759;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 KEIKGRTV 150  
DB 177 KEIKGRTV 184

RESULT 6  
US-08-676-974-1  
Sequence 1, Application US/08676974  
Patent No. 5770422  
GENERAL INFORMATION:  
APPLICANT: COLLINS, KATHLEEN  
TITLE OF INVENTION: Human Telomerase  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Science & Technology Law Group  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/676,974  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UCB96-055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415)343-4342  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 759 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-676-974-1

Query Match 4.4%; Score 8; DB 1; Length 759;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 KEIKGRTV 150  
DB 177 KEIKGRTV 184



Db 177 KEIKGRTV 184

RESULT 7

US-09-098-487-1

; Sequence 1, Application US/09098487

; Patent No. 5917025

; GENERAL INFORMATION:

; APPLICANT: COLLINS, Kathleen

; TITLE OF INVENTION: Human Telomerase

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Science & Technology Law Group

; STREET: 268 Bush Street, Suite 3200

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/098,487

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman Ph.D., Richard A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UCB96-055

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415)343-4341

; TELEFAX: (415)343-4342

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 759 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide

US-09-098-487-1

Query Match 4.4%; Score 8; DB 2; Length 759;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 KEIKGRTV 150

Db 177 KEIKGRTV 184

RESULT 8

US-09-428-082B-1048

; Sequence 1048, Application US/09428082B

; Patent No. 6660843

; GENERAL INFORMATION:

; APPLICANT: FEIGE, ULRICH

; APPLICANT: LIU, CHUAN-FA

; APPLICANT: CHEETHAM, JANET C.

; APPLICANT: BOONE, THOMAS CHARLES

; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS

; FILE REFERENCE: A-527

; CURRENT APPLICATION NUMBER: US/09/428,082B

; CURRENT FILING DATE: 1999-10-22

; PRIOR APPLICATION NUMBER: 60/105,371

; PRIOR FILING DATE: 1998-10-23

; NUMBER OF SEQ ID NOS: 1133

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1048

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Artificial Sequence

Qy 177 KEIKGRTV 184

Db 177 KEIKGRTV 184

FEATURE:

OTHER INFORMATION: IL-1 ANTAGONIST

FEATURE:

NAME/KEY: misc feature

LOCATION: (20)..(20)

OTHER INFORMATION: Fc domain attached at Position 20 of the C-terminus

US-09-428-082B-1048

Query Match 3.9%; Score 7; DB 4; Length 20;

Best Local Similarity 100.0%; Pred. No. 6.6;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 LPLGGGG 170

Db 13 LPLGGGG 19

RESULT 9

US-09-107-532A-6096

; Sequence 6096, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Arinello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 6096:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 133 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...133

; SEQUENCE DESCRIPTION: SEQ ID NO: 6096:

US-09-107-532A-6096

Query Match 3.9%; Score 7; DB 4; Length 133;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 TDLKGLP 67



```
Db      105 TDLKGLP 111
|||||
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21201
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21201

Query Match      3.9%; Score 7; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      167 GGGGARI 173
      |||||
Db      57 GGGGARI 63

RESULT 13
US-09-134-000C-3711
; Sequence 3711, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3711
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3711

Query Match      3.9%; Score 7; DB 4; Length 157;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      61 TDLKGLP 67
      |||||
Db      128 TDLKGLP 134

RESULT 14
US-09-543-681A-4466
; Sequence 4466, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4466
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4466

Query Match      3.9%; Score 7; DB 4; Length 178;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      155 GGDNYSD 161
      |||||
Db      154 GGDNYSD 160
```

---

```
Db      105 TDLKGLP 111
|||||
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21201
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21201

Query Match      3.9%; Score 7; DB 4; Length 148;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      132 KETLLAP 138
      |||||
Db      70 KETLLAP 76

RESULT 11
US-09-270-767-53140
; Sequence 53140, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53140
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53140

Query Match      3.9%; Score 7; DB 4; Length 148;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      132 KETLLAP 138
      |||||
Db      70 KETLLAP 76

RESULT 12
US-09-252-991A-21201
; Sequence 21201, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
```

```

RESULT 15
US-09-270-767-37351
; Sequence 37351, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37351
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-37351

Query Match          3.9%; Score 7; DB 4; Length 191;
Best Local Similarity 100.0%; Pred.No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps

QY      118 GDLPLRV 124
Db       35 GDLPLRV 41
|||||||

RESULT 16
US-09-270-767-52568
; Sequence 52568, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52568
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-52568

Query Match          3.9%; Score 7; DB 4; Length 191;
Best Local Similarity 100.0%; Pred.No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps

QY      118 GDLPLRV 124
Db       35 GDLPLRV 41
|||||||

RESULT 17
US-09-428-082B-1062
; Sequence 1062, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B
; CURRENT FILING DATE: 1999-10-22

```

```
Query Match          3.9%; Score 7; DB 4; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GGGARIA 174
    |||||
Db 115 GGGARIA 121

RESULT 20
US-09-252-991A-22291
; Sequence 22291, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22291
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22291

Query Match          3.9%; Score 7; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 AGLOAHG 98
    |||||
Db 153 AGLOAHG 159

RESULT 21
US-09-489-039A-11461
; Sequence 11461, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11461
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11461

Query Match          3.9%; Score 7; DB 4; Length 535;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 PRLVVKA 127
    |||||
Db 52 PRLVVKA 58

RESULT 22
US-09-248-796A-20375
; Sequence 20375, Application US/09248796A
```

```
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20375
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20375

Query Match          3.9%; Score 7; DB 4; Length 571;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GHLTAGL 94
    |||||
Db 232 GHLTAGL 238

RESULT 23
US-09-270-767-46196
; Sequence 46196, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46196
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46196

Query Match          3.9%; Score 7; DB 4; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGA 171
    |||||
Db 216 PLGGGGA 222

RESULT 24
US-09-270-767-41618
; Sequence 41618, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41618
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
```

US-09-270-767-41618

Query Match 3.9%; Score 7; DB 4; Length 674;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 165 PLGGGGA 171  
Db 285 PLGGGGA 291

RESULT 25

US-09-270-767-44535  
; Sequence 44535, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 44535  
; LENGTH: 679  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-44535

Query Match 3.9%; Score 7; DB 4; Length 679;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 165 PLGGGGA 171  
Db 290 PLGGGGA 296

RESULT 26

US-09-486-072-1  
; Sequence 1, Application US/09486072  
; Patent No. 6489155  
; GENERAL INFORMATION:  
; APPLICANT: Masanori TAKAYAMA, et al.  
; TITLE OF INVENTION: GENES  
; FILE REFERENCE: 11202/1  
; CURRENT APPLICATION NUMBER: US/09/486,072  
; CURRENT FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: PCT/JP98/02310  
; PRIOR FILING DATE: 1998-05-26  
; PRIOR APPLICATION NUMBER: JP252624/97  
; PRIOR FILING DATE: 1997-09-03  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 814  
; TYPE: PRT  
; ORGANISM: Bacteria  
US-09-486-072-1

Query Match 3.9%; Score 7; DB 4; Length 814;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 LGGGGAR 172  
Db 361 LGGGGAR 367

RESULT 27

US-09-604-978-11  
; Sequence 11, Application US/09604978  
; Patent No. 6455674

; GENERAL INFORMATION:  
; APPLICANT: Einat, Paz  
; TITLE OF INVENTION: HYPOXIA-REGULATED GENES  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: KOHN & ASSOCIATES  
; STREET: 30500 No. 6455674thwestern Hwy., Suite 401  
; CITY: Farmington Hills  
; STATE: Michigan  
; COUNTRY: U.S.  
; ZIP: 48334  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/604,978  
; FILING DATE: 28-Jun-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/138,112  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kohn, Kenneth I.  
; REGISTRATION NUMBER: 30,955  
; REFERENCE/DOCKET NUMBER: 0168.00034  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (248) 539-5050  
; TELEFAX: (248) 5395055  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 864 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-604-978-11

Query Match 3.9%; Score 7; DB 4; Length 864;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLTSITS 22  
Db 650 LLTSITS 656

RESULT 28  
US-09-604-728-11  
; Sequence 11, Application US/09604728  
; Patent No. 6555667  
; GENERAL INFORMATION:  
; APPLICANT: Einat, Paz  
; TITLE OF INVENTION: HYPOXIA-REGULATED GENES  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: KOHN & ASSOCIATES  
; STREET: 30500 No. 6555667thwestern Hwy., Suite 401  
; CITY: Farmington Hills  
; STATE: Michigan  
; COUNTRY: U.S.  
; ZIP: 48334  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/604,728  
FILING DATE: 28-Jun-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/138,112  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohn, Kenneth I.  
REGISTRATION NUMBER: 30,955  
REFERENCE/DOCKET NUMBER: 0168.00034  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (248) 539-5050  
TELEFAX: (248) 5395055  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 864 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-604-728-11  
Query Match 3.9%; Score 7; DB 4; Length 864;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
QY 16 LITSITS 22  
DB 650 LITSITS 656  
RESULT 29  
US-10-325-878-11  
Sequence 11, Application US/10325878  
Patent No. 6740738  
GENERAL INFORMATION:  
APPLICANT: Binat, Paz  
Skaliter, Rami  
TITLE OF INVENTION: HYPOXIA-REGULATED GENES  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KOHN & ASSOCIATES  
STREET: 30500 No. 6740738thwestern Hwy., Suite 401  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/325,878  
FILING DATE: 23-Dec-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/138,112  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohn, Kenneth I.  
REGISTRATION NUMBER: 30,955  
REFERENCE/DOCKET NUMBER: 0168.00034  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (248) 539-5050  
TELEFAX: (248) 5395055  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 864 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-10-325-878-11  
Query Match 3.9%; Score 7; DB 4; Length 864;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02; Indels 0;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
QY 16 LITSITS 22  
DB 650 LITSITS 656  
RESULT 30  
US-09-180-245-2  
Sequence 2, Application US/09180245  
Patent No. 6602659  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A  
TITLE OF INVENTION: Methods of and Kits and Compositions for Diagnosing  
TITLE OF INVENTION: Colorectal Tumors and Metastasis Thereof  
FILE REFERENCE: TUJ2161  
CURRENT APPLICATION NUMBER: US/09/180,245  
CURRENT FILING DATE: 1999-03-11  
EARLIER APPLICATION NUMBER: PCT/US97/07467  
EARLIER FILING DATE: 1997-05-02  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 1073  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Carboxy terminal tail - nucleotides 3148-3336,  
OTHER INFORMATION: corresponding to amino acids 1011-1073.  
US-09-180-245-2  
Query Match 3.9%; Score 7; DB 4; Length 1073;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 120 LPRLVVK 126  
DB 801 LPRLVVK 807  
RESULT 31  
US-09-819-249-2  
Sequence 2, Application US/09819249  
Patent No. 6767704  
GENERAL INFORMATION:  
APPLICANT: Waldman, Scott A.  
APPLICANT: Park, Jason  
APPLICANT: Schulz, Stephanie  
TITLE OF INVENTION: Compositions And Methods For Identifying And Targeting Cancer Cells  
TITLE OF INVENTION: Alimentary Canal Origin  
FILE REFERENCE: TUJ2412  
CURRENT APPLICATION NUMBER: US/09/819,249  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: 60/192,229  
PRIOR FILING DATE: 2000-03-27  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 1073  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-819-249-2  
Query Match 3.9%; Score 7; DB 4; Length 1073;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 LPLRVK 126  
| | | | |  
Db 801 LPLRVK 807

RESULT 32  
US-07-623-033-2  
; Sequence 2, Application US/07623033  
; Patent No. 5237051  
; GENERAL INFORMATION:  
; APPLICANT: Garbers, David L.  
; APPLICANT: Schulz, Stephanie  
; TITLE OF INVENTION: CLONING THE ENTEROTOXIN RECEPTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT  
; STREET: 100 South Wacker Drive - Suite 960  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60606-4002  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/623,033  
; FILING DATE: 19901206  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kohn, Kenneth I.  
; REGISTRATION NUMBER: 30,955  
; REFERENCE/DOCKET NUMBER: VU9018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 456-8000  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1075 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-623-033-2

Query Match 3.9%; Score 7; DB 1; Length 1075;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 LPLRVK 126  
| | | | |  
Db 803 LPLRVK 809

RESULT 33  
5245013-14  
; Patent No. 5245013  
; APPLICANT: Ulevitch, Richard; Tobias, Peter  
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC  
; ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES  
; NUMBER OF SEQUENCES: 20  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/367,454  
; FILING DATE: 01-JUN-1989  
; APPLICATION NUMBER: 6,710  
; FILING DATE: 30-DEC-1986  
; APPLICATION NUMBER: 728,833  
; FILING DATE: 30-APR-1985  
; SEQ ID NO: 14:  
; LENGTH: 14  
5245013-14

Query Match 3.3%; Score 6; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 TDKGLQ 58  
| | | | |  
Db 9 TDKGLQ 14

RESULT 34  
5245013-15  
; Patent No. 5245013  
; APPLICANT: Ulevitch, Richard; Tobias, Peter  
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC  
; ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES  
; NUMBER OF SEQUENCES: 20  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/367,454  
; FILING DATE: 01-JUN-1989  
; APPLICATION NUMBER: 6,710  
; FILING DATE: 30-DEC-1986  
; APPLICATION NUMBER: 728,833  
; FILING DATE: 30-APR-1985  
; SEQ ID NO: 15:  
; LENGTH: 16  
5245013-15

Query Match 3.3%; Score 6; DB 6; Length 16;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 TDKGLQ 58  
| | | | |  
Db 7 TDKGLQ 12

RESULT 35  
US-09-077-354B-5  
; Sequence 5, Application US/09077354B  
; Patent No. 6255096  
; GENERAL INFORMATION:  
; APPLICANT: HOPWOOD, JOHN JOSEPH; SCOTT, HAMISH STEELE;  
; APPLICANT: WEBER, BIRGIT; BLANCH, LIANNE; ANSON, DONALD STEWART  
; TITLE OF INVENTION: SYNTHETIC MAMMALIAN  
; MOLECULE TYPE: N-ACETYLGLUCOSAMINIDASE AND GENETIC SEQUENCES ENCODING SAME  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 GARDEN CITY PLAZA  
; CITY: GARDEN CITY  
; STATE: NEW YORK  
; COUNTRY: UNITED STATES  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/077,354B  
; FILING DATE: 22-APRIL-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/00747  
; FILING DATE: 22-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PORALSKY, ANN R.  
; REGISTRATION NUMBER: 34,697  
; REFERENCE/DOCKET NUMBER: 12416  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516 742 4343  
; TELEFAX: 516 742 4366  
; INFORMATION FOR SEQ ID NO: 5:

Tue Nov 2 12:20:07 2004

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Modified-site, glycosylated or
; NAME/KEY: phosphorylated, wherein Xaa may be any
; NAME/KEY: amino acid residue, preferably Arg.
; LOCATION: 16
;
US-09-077-354B-5
Query Match 3.3%; Score 6; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGA 171
Db 9 LGGGGA 14

RESULT 36
5245013-11
; APPLICANT: Ulevitch, Richard; Tobias, Peter
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC
; ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/367,454
; FILING DATE: 01-JUN-1989
; APPLICATION NUMBER: 6,710
; FILING DATE: 30-DEC-1986
; APPLICATION NUMBER: 728,833
; FILING DATE: 30-APR-1985
; SEQ ID NO:11:
; LENGTH: 20
5245013-11
Query Match 3.3%; Score 6; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TDKGLQ 58
Db 10 TDKGLQ 15

RESULT 37
US-09-383-062-54
; Sequence 54, Application US/09383062
; Patent No. 6670194
; GENERAL INFORMATION:
; APPLICANT: Aebersold, Rudolf H.
; APPLICANT: Gelb, Michael H
; APPLICANT: Gygi, Steven
; APPLICANT: Scott, C R
; APPLICANT: Turecek, Frantisek
; APPLICANT: Gerber, Scott A
; APPLICANT: Rist, Beate
; TITLE OF INVENTION: Rapid Quantitative Analysis of Proteins or Protein
; FILE REFERENCE: Function in Complex Mixture
; CURRENT FILING DATE: 1999-08-25
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 54
; LENGTH: 25
; TYPE: PRT
; ORGANISM: yeast
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (14)
; OTHER INFORMATION: C at position 14 is ICAT-labeled cysteinyl
; OTHER INFORMATION: residue.
US-09-383-062-54
Query Match 3.3%; Score 6; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GFHIHE 77
Db 1 GFHIHE 6

RESULT 38
US-09-270-767-34704
; Sequence 34704, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34704
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-34704
Query Match 3.3%; Score 6; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGA 171
Db 8 LGGGGA 13

RESULT 39
US-09-270-767-49921
; Sequence 49921, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49921
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-49921
Query Match 3.3%; Score 6; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGA 171
Db 8 LGGGGA 13
```

```
RESULT 40
US-08-023-980B-27
; Sequence 27, Application US/08023980B
; Patent No. 5843641
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Rosen, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: TREATMENT AND PREVENTION OF DISEASES OF CELL DEATH
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 585 Commercial Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-1024
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/023,980B
; FILING DATE: 26-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/177001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/723-4123
; TELEFAX: 617/723-8962
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-023-980B-27

Query Match 3.3%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 GEGGFH 74
Db 18 GEGGFH 23

Search completed: October 26, 2004, 10:04:25
Job time : 53 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 10:03:31 ; Search time 128 Seconds

(without alignments)  
455.286 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 180

Sequence: 1 MKIKLFFVTSIVTISLLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1364641 seqs, 323758627 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description          |
|------------|-------|---------------|--------|----|----------------------|
| 1          | 8     | 4.4           | 18     | 14 | US-10-225-567A-1150  |
| 2          | 8     | 4.4           | 56     | 15 | US-10-424-599-250991 |
| 3          | 8     | 4.4           | 92     | 16 | US-10-437-963-124448 |
| 4          | 8     | 4.4           | 201    | 16 | US-10-437-963-117523 |
| 5          | 8     | 4.4           | 244    | 15 | US-10-425-114-39062  |
| 6          | 8     | 4.4           | 369    | 16 | US-10-437-963-164580 |
| 7          | 8     | 4.4           | 476    | 14 | US-10-369-493-2997   |
| 8          | 8     | 4.4           | 532    | 10 | US-09-826-509-521    |
| 9          | 8     | 4.4           | 532    | 14 | US-10-225-567A-195   |
| 10         | 8     | 4.4           | 532    | 14 | US-10-366-288-34     |
| 11         | 8     | 4.4           | 532    | 14 | US-10-280-066-48     |
| 12         | 7     | 3.9           | 20     | 14 | US-10-609-217-1048   |
| 13         | 7     | 3.9           | 20     | 15 | US-10-609-217-1048   |

Sequence 1048, Ap  
Sequence 1048, Ap  
Sequence 1048, Ap  
Sequence 1048, Ap  
Sequence 1048, Ap  
Sequence 163985, A  
Sequence 50522, A  
Sequence 140344, A  
Sequence 140344, A  
Sequence 153655, A  
Sequence 158607, A  
Sequence 40171, A  
Sequence 261961, A  
Sequence 168698, A  
Sequence 160841, A  
Sequence 190020, A  
Sequence 57172, A  
Sequence 57138, A  
Sequence 40863, A  
Sequence 58300, A  
Sequence 144207, A  
Sequence 37046, A  
Sequence 170400, A  
Sequence 40660, A  
Sequence 167300, A  
Sequence 269534, A  
Sequence 2908, Ap  
Sequence 268560, A  
Sequence 42635, A  
Sequence 111546, A  
Sequence 194937, A  
Sequence 60, Appl  
Sequence 194934, A  
Sequence 71299, A  
Sequence 145553, A  
Sequence 178478, A  
Sequence 186201, A  
Sequence 51313, A  
Sequence 1062, Ap  
Sequence 1062, Ap  
Sequence 1062, Ap  
Sequence 1062, Ap  
Sequence 1062, Ap  
Sequence 273159, A  
Sequence 174632, A  
Sequence 178, App  
Sequence 1658, Ap  
Sequence 1679, Ap  
Sequence 173885, A  
Sequence 174633, A  
Sequence 6834, A  
Sequence 171705, A  
Sequence 68424, A  
Sequence 159211, A  
Sequence 16479, A  
Sequence 46306, A  
Sequence 46004, A  
Sequence 45741, A  
Sequence 135287, A  
Sequence 1789, Ap  
Sequence 3204, Ap  
Sequence 184096, A  
Sequence 172356, A  
Sequence 103746, A  
Sequence 3257, Ap  
Sequence 2262, Ap  
Sequence 173, App  
Sequence 24231, A  
Sequence 58924, A  
Sequence 155587, A  
Sequence 11, Appl  
Sequence 11, Appl

7 3.9 20 15 US-10-632-388-1048  
7 3.9 20 15 US-10-651-723-1048  
7 3.9 20 15 US-10-645-761-1048  
7 3.9 20 15 US-10-666-696-1048  
7 3.9 20 15 US-10-653-048-1048  
7 3.9 53 16 US-10-437-963-163985  
7 3.9 69 16 US-10-767-701-50522  
7 3.9 85 16 US-10-437-963-140344  
7 3.9 96 15 US-10-425-114-53565  
7 3.9 105 15 US-10-424-599-158607  
7 3.9 107 15 US-10-425-114-40171  
7 3.9 108 15 US-10-424-599-261961  
7 3.9 109 16 US-10-437-963-168698  
7 3.9 118 15 US-10-424-599-160841  
7 3.9 121 16 US-10-437-963-190020  
7 3.9 123 15 US-10-282-122A-57138  
7 3.9 124 15 US-10-282-122A-57138  
7 3.9 124 16 US-10-767-701-40863  
7 3.9 140 16 US-10-767-701-58300  
7 3.9 141 16 US-10-437-963-144207  
7 3.9 141 16 US-10-767-701-37046  
7 3.9 142 16 US-10-437-963-170400  
7 3.9 143 16 US-10-767-701-40660  
7 3.9 144 16 US-10-437-963-167300  
7 3.9 150 15 US-10-424-599-269534  
7 3.9 157 14 US-10-104-047-2908  
7 3.9 166 15 US-10-424-599-268560  
7 3.9 174 16 US-10-767-701-42635  
7 3.9 177 16 US-10-437-963-111546  
7 3.9 183 15 US-10-424-599-194937  
7 3.9 186 14 US-10-320-800-60  
7 3.9 188 15 US-10-424-599-194934  
7 3.9 188 15 US-10-425-114-71299  
7 3.9 192 16 US-10-437-963-145553  
7 3.9 192 16 US-10-437-963-178478  
7 3.9 210 16 US-10-437-963-186201  
7 3.9 233 15 US-10-282-122A-51313  
7 3.9 248 15 US-10-609-217-1062  
7 3.9 248 15 US-10-632-388-1062  
7 3.9 248 15 US-10-651-723-1062  
7 3.9 248 15 US-10-645-761-1062  
7 3.9 248 15 US-10-666-696-1062  
7 3.9 260 15 US-10-653-048-1062  
7 3.9 285 16 US-10-424-599-273159  
7 3.9 289 15 US-10-437-963-174632  
7 3.9 289 15 US-10-074-978A-178  
7 3.9 303 15 US-10-374-780A-1658  
7 3.9 303 15 US-10-412-699B-1679  
7 3.9 303 16 US-10-437-963-173885  
7 3.9 307 16 US-10-437-963-174633  
7 3.9 337 15 US-10-425-114-66834  
7 3.9 351 16 US-10-437-963-171705  
7 3.9 357 15 US-10-425-114-68424  
7 3.9 378 16 US-10-437-963-159211  
7 3.9 398 14 US-10-369-493-16479  
7 3.9 405 15 US-10-282-122A-46306  
7 3.9 409 15 US-10-282-122A-46004  
7 3.9 410 16 US-10-767-701-46545  
7 3.9 428 15 US-10-425-114-45741  
7 3.9 447 16 US-10-437-963-135287  
7 3.9 455 14 US-10-369-493-1789  
7 3.9 456 15 US-10-320-797-3204  
7 3.9 458 16 US-10-437-963-184096  
7 3.9 535 16 US-10-437-963-172356  
7 3.9 540 16 US-10-437-963-103746  
7 3.9 601 15 US-10-320-797-3257  
7 3.9 702 14 US-10-094-749-2262  
7 3.9 702 15 US-10-074-978A-173  
7 3.9 727 15 US-10-424-599-246231  
7 3.9 764 15 US-10-425-114-58924  
7 3.9 813 16 US-10-437-963-155587  
7 3.9 864 14 US-10-091-333-11  
7 3.9 864 14 US-10-325-878-11

|     |   |     |      |    |                       |                    |     |   |     |    |                      |                    |
|-----|---|-----|------|----|-----------------------|--------------------|-----|---|-----|----|----------------------|--------------------|
| 87  | 7 | 3.9 | 871  | 15 | US-10-320-797-3313    | Sequence 3313, Ap  | 160 | 6 | 3.3 | 15 | US-10-424-599-170939 | Sequence 170939, A |
| 88  | 7 | 3.9 | 1048 | 15 | US-10-282-1222A-64852 | Sequence 64852, A  | 161 | 6 | 3.3 | 16 | US-10-167-701-56417  | Sequence 56417, A  |
| 89  | 7 | 3.9 | 1073 | 14 | US-09-819-249-2       | Sequence 2, Appli  | 162 | 6 | 3.3 | 16 | US-10-767-701-58020  | Sequence 58020, A  |
| 90  | 7 | 3.9 | 1073 | 14 | US-10-157-031-18      | Sequence 18, Appl  | 163 | 6 | 3.3 | 15 | US-10-424-599-177954 | Sequence 177954, A |
| 91  | 7 | 3.9 | 1073 | 16 | US-10-695-578-2       | Sequence 2, Appli  | 164 | 6 | 3.3 | 67 | US-09-986-480-425    | Sequence 425, App  |
| 92  | 7 | 3.9 | 1119 | 16 | US-10-437-963-145554  | Sequence 145554, A | 165 | 6 | 3.3 | 67 | US-10-424-599-262497 | Sequence 262497, A |
| 93  | 7 | 3.9 | 1232 | 16 | US-10-282-1222A-52194 | Sequence 52194, A  | 166 | 6 | 3.3 | 67 | US-10-767-701-48436  | Sequence 48436, A  |
| 94  | 7 | 3.9 | 1324 | 16 | US-10-437-963-182767  | Sequence 182767, A | 167 | 6 | 3.3 | 68 | US-10-424-599-158466 | Sequence 158466, A |
| 95  | 7 | 3.9 | 1327 | 15 | US-10-074-978A-176    | Sequence 176, App  | 168 | 6 | 3.3 | 69 | US-10-424-599-205028 | Sequence 205028, A |
| 96  | 7 | 3.9 | 1327 | 15 | US-10-074-978A-176    | Sequence 176, App  | 169 | 6 | 3.3 | 69 | US-10-424-599-257072 | Sequence 257072, A |
| 97  | 7 | 3.9 | 1389 | 15 | US-10-074-978A-175    | Sequence 175, App  | 170 | 6 | 3.3 | 70 | US-10-074-475-203    | Sequence 203, App  |
| 98  | 7 | 3.9 | 1425 | 15 | US-10-074-978A-174    | Sequence 174, App  | 171 | 6 | 3.3 | 71 | US-10-424-599-176190 | Sequence 176190, A |
| 99  | 7 | 3.9 | 1487 | 15 | US-10-282-1222A-77442 | Sequence 77442, A  | 172 | 6 | 3.3 | 72 | US-10-437-963-146566 | Sequence 146566, A |
| 100 | 7 | 3.9 | 1605 | 16 | US-10-437-963-171864  | Sequence 171864, A | 173 | 6 | 3.3 | 72 | US-10-437-963-146566 | Sequence 146566, A |
| 101 | 7 | 3.9 | 1849 | 16 | US-10-437-963-110584  | Sequence 110584, A | 174 | 6 | 3.3 | 73 | US-09-864-761-38999  | Sequence 38999, A  |
| 102 | 7 | 3.9 | 1858 | 16 | US-10-437-963-171853  | Sequence 171853, A | 175 | 6 | 3.3 | 73 | US-10-424-599-155186 | Sequence 155186, A |
| 103 | 7 | 3.9 | 1925 | 16 | US-10-437-963-152541  | Sequence 152541, A | 176 | 6 | 3.3 | 73 | US-10-437-963-128020 | Sequence 128020, A |
| 104 | 7 | 3.9 | 2098 | 16 | US-10-437-963-110559  | Sequence 110559, A | 177 | 6 | 3.3 | 73 | US-10-437-963-168183 | Sequence 168183, A |
| 105 | 6 | 3.3 | 9    | 14 | US-10-144-188-58      | Sequence 110559, A | 178 | 6 | 3.3 | 74 | US-10-424-599-243900 | Sequence 243900, A |
| 106 | 6 | 3.3 | 13   | 14 | US-10-285-394-389     | Sequence 58, Appl  | 179 | 6 | 3.3 | 74 | US-10-437-963-147720 | Sequence 147720, A |
| 107 | 6 | 3.3 | 13   | 15 | US-10-432-422-131     | Sequence 389, App  | 180 | 6 | 3.3 | 76 | US-09-990-046-4      | Sequence 4, Appli  |
| 108 | 6 | 3.3 | 18   | 10 | US-09-836-613-5       | Sequence 131, App  | 181 | 6 | 3.3 | 76 | US-10-029-386-33871  | Sequence 33871, A  |
| 109 | 6 | 3.3 | 20   | 9  | US-09-813-333-78      | Sequence 5, Appli  | 182 | 6 | 3.3 | 76 | US-10-424-599-238879 | Sequence 238879, A |
| 110 | 6 | 3.3 | 20   | 13 | US-10-044-703-78      | Sequence 78, Appl  | 183 | 6 | 3.3 | 76 | US-10-437-963-187695 | Sequence 187695, A |
| 111 | 6 | 3.3 | 20   | 15 | US-10-239-103-78      | Sequence 78, Appl  | 184 | 6 | 3.3 | 77 | US-10-424-599-259260 | Sequence 259260, A |
| 112 | 6 | 3.3 | 21   | 9  | US-09-864-761-40560   | Sequence 40560, A  | 185 | 6 | 3.3 | 77 | US-10-424-599-262009 | Sequence 262009, A |
| 113 | 6 | 3.3 | 24   | 14 | US-10-145-206-12      | Sequence 12, Appl  | 186 | 6 | 3.3 | 78 | US-10-424-599-232759 | Sequence 232759, A |
| 114 | 6 | 3.3 | 25   | 9  | US-09-839-884-54      | Sequence 54, Appl  | 187 | 6 | 3.3 | 79 | US-10-291-265-842    | Sequence 842, App  |
| 115 | 6 | 3.3 | 25   | 10 | US-09-839-884-54      | Sequence 54, Appl  | 188 | 6 | 3.3 | 79 | US-10-424-599-250797 | Sequence 250797, A |
| 116 | 6 | 3.3 | 26   | 9  | US-09-864-761-35466   | Sequence 35466, A  | 189 | 6 | 3.3 | 80 | US-10-335-977-5710   | Sequence 5710, Ap  |
| 117 | 6 | 3.3 | 26   | 9  | US-09-864-761-36004   | Sequence 36004, A  | 190 | 6 | 3.3 | 80 | US-10-335-977-5709   | Sequence 5709, Ap  |
| 118 | 6 | 3.3 | 32   | 14 | US-10-174-410-281     | Sequence 281, App  | 191 | 6 | 3.3 | 80 | US-10-437-963-103847 | Sequence 103847, A |
| 119 | 6 | 3.3 | 34   | 10 | US-09-986-480-426     | Sequence 426, App  | 192 | 6 | 3.3 | 80 | US-10-437-963-150647 | Sequence 150647, A |
| 120 | 6 | 3.3 | 38   | 15 | US-10-371-073-23      | Sequence 23, Appl  | 193 | 6 | 3.3 | 81 | US-10-437-963-204221 | Sequence 204221, A |
| 121 | 6 | 3.3 | 39   | 9  | US-09-864-761-49109   | Sequence 49109, A  | 194 | 6 | 3.3 | 81 | US-10-424-599-270119 | Sequence 270119, A |
| 122 | 6 | 3.3 | 39   | 14 | US-10-029-386-311364  | Sequence 31364, A  | 195 | 6 | 3.3 | 82 | US-10-437-963-167739 | Sequence 167739, A |
| 123 | 6 | 3.3 | 43   | 10 | US-09-892-877-389     | Sequence 389, App  | 196 | 6 | 3.3 | 83 | US-10-424-599-187092 | Sequence 187092, A |
| 124 | 6 | 3.3 | 43   | 10 | US-09-948-783-389     | Sequence 389, App  | 197 | 6 | 3.3 | 84 | US-10-437-963-135954 | Sequence 135954, A |
| 125 | 6 | 3.3 | 43   | 15 | US-10-424-599-256077  | Sequence 256077, A | 198 | 6 | 3.3 | 84 | US-10-437-963-169092 | Sequence 169092, A |
| 126 | 6 | 3.3 | 44   | 15 | US-10-424-599-163693  | Sequence 163693, A | 199 | 6 | 3.3 | 85 | US-10-424-599-159783 | Sequence 159783, A |
| 127 | 6 | 3.3 | 45   | 15 | US-10-424-599-188241  | Sequence 188241, A | 200 | 6 | 3.3 | 85 | US-10-437-963-113725 | Sequence 113725, A |
| 128 | 6 | 3.3 | 45   | 16 | US-10-767-701-62107   | Sequence 62107, A  | 201 | 6 | 3.3 | 86 | US-10-437-963-161950 | Sequence 161950, A |
| 129 | 6 | 3.3 | 47   | 15 | US-10-424-599-239062  | Sequence 239062, A | 202 | 6 | 3.3 | 86 | US-09-864-761-47995  | Sequence 47995, A  |
| 130 | 6 | 3.3 | 47   | 16 | US-10-437-963-120410  | Sequence 120410, A | 203 | 6 | 3.3 | 86 | US-10-424-599-207725 | Sequence 207725, A |
| 131 | 6 | 3.3 | 48   | 9  | US-09-864-761-38338   | Sequence 38338, A  | 204 | 6 | 3.3 | 86 | US-10-437-963-132923 | Sequence 132923, A |
| 132 | 6 | 3.3 | 48   | 16 | US-10-437-963-133014  | Sequence 133014, A | 205 | 6 | 3.3 | 86 | US-10-437-963-136907 | Sequence 136907, A |
| 133 | 6 | 3.3 | 48   | 16 | US-10-437-963-179723  | Sequence 179723, A | 206 | 6 | 3.3 | 87 | US-10-001-843-153    | Sequence 153, App  |
| 134 | 6 | 3.3 | 50   | 15 | US-10-424-599-276411  | Sequence 276411, A | 207 | 6 | 3.3 | 87 | US-10-156-761-11193  | Sequence 11193, A  |
| 135 | 6 | 3.3 | 51   | 15 | US-10-424-599-179703  | Sequence 179703, A | 208 | 6 | 3.3 | 88 | US-10-437-963-177905 | Sequence 177905, A |
| 136 | 6 | 3.3 | 51   | 16 | US-10-437-963-106496  | Sequence 106496, A | 209 | 6 | 3.3 | 89 | US-10-437-963-179091 | Sequence 179091, A |
| 137 | 6 | 3.3 | 52   | 9  | US-09-925-299-1519    | Sequence 1519, Ap  | 210 | 6 | 3.3 | 89 | US-10-424-599-175481 | Sequence 175481, A |
| 138 | 6 | 3.3 | 52   | 9  | US-09-905-176-4       | Sequence 4, Appli  | 211 | 6 | 3.3 | 90 | US-10-767-701-42059  | Sequence 42059, A  |
| 139 | 6 | 3.3 | 52   | 10 | US-09-925-299-1519    | Sequence 1519, Ap  | 212 | 6 | 3.3 | 90 | US-10-424-599-202015 | Sequence 202015, A |
| 140 | 6 | 3.3 | 53   | 9  | US-09-905-176-3       | Sequence 3, Appli  | 213 | 6 | 3.3 | 90 | US-10-437-963-174330 | Sequence 174330, A |
| 141 | 6 | 3.3 | 54   | 16 | US-10-437-963-104694  | Sequence 104694, A | 214 | 6 | 3.3 | 91 | US-10-437-963-185511 | Sequence 185511, A |
| 142 | 6 | 3.3 | 56   | 15 | US-10-424-599-173935  | Sequence 173935, A | 215 | 6 | 3.3 | 91 | US-10-424-599-184624 | Sequence 184624, A |
| 143 | 6 | 3.3 | 56   | 16 | US-10-424-599-180278  | Sequence 180278, A | 216 | 6 | 3.3 | 92 | US-10-425-114-61847  | Sequence 61847, A  |
| 144 | 6 | 3.3 | 57   | 14 | US-10-029-386-28186   | Sequence 28186, A  | 217 | 6 | 3.3 | 92 | US-10-437-963-177496 | Sequence 177496, A |
| 145 | 6 | 3.3 | 57   | 15 | US-10-424-599-260447  | Sequence 260447, A | 218 | 6 | 3.3 | 93 | US-10-002-3448-256   | Sequence 256, App  |
| 146 | 6 | 3.3 | 57   | 15 | US-10-424-599-278897  | Sequence 278897, A | 219 | 6 | 3.3 | 93 | US-10-419-296-12     | Sequence 29906, A  |
| 147 | 6 | 3.3 | 58   | 15 | US-10-425-114-45121   | Sequence 45121, A  | 220 | 6 | 3.3 | 93 | US-10-424-599-239906 | Sequence 239906, A |
| 148 | 6 | 3.3 | 58   | 15 | US-10-767-701-31893   | Sequence 31893, A  | 221 | 6 | 3.3 | 93 | US-10-437-963-151745 | Sequence 151745, A |
| 149 | 6 | 3.3 | 59   | 15 | US-10-424-599-219197  | Sequence 219197, A | 222 | 6 | 3.3 | 93 | US-10-437-963-163002 | Sequence 163002, A |
| 150 | 6 | 3.3 | 59   | 15 | US-10-424-599-272278  | Sequence 272278, A | 223 | 6 | 3.3 | 94 | US-10-424-599-227050 | Sequence 227050, A |
| 151 | 6 | 3.3 | 59   | 16 | US-10-437-963-150712  | Sequence 150712, A | 224 | 6 | 3.3 | 94 | US-10-437-963-134836 | Sequence 134836, A |
| 152 | 6 | 3.3 | 61   | 11 | US-10-424-599-154711  | Sequence 154711, A | 225 | 6 | 3.3 | 95 | US-10-128-714-3550   | Sequence 3550, Ap  |
| 153 | 6 | 3.3 | 61   | 11 | US-09-864-408A-9046   | Sequence 9046, Ap  | 226 | 6 | 3.3 | 95 | US-10-424-599-230356 | Sequence 230356, A |
| 154 | 6 | 3.3 | 62   | 15 | US-10-424-599-173817  | Sequence 173817, A | 227 | 6 | 3.3 | 96 | US-10-424-599-216852 | Sequence 216852, A |
| 155 | 6 | 3.3 | 62   | 15 | US-10-424-599-242159  | Sequence 242159, A | 228 | 6 | 3.3 | 96 | US-10-437-963-166457 | Sequence 166457, A |
| 156 | 6 | 3.3 | 62   | 16 | US-10-437-963-130092  | Sequence 130092, A | 229 | 6 | 3.3 | 96 | US-10-437-963-170185 | Sequence 170185, A |
| 157 | 6 | 3.3 | 63   | 15 | US-10-424-599-195550  | Sequence 195550, A | 230 | 6 | 3.3 | 96 |                      |                    |
| 158 | 6 | 3.3 | 63   | 16 | US-10-437-963-204001  | Sequence 204001, A | 231 | 6 | 3.3 | 96 |                      |                    |
| 159 | 6 | 3.3 | 63   | 16 | US-10-767-701-48356   | Sequence 48356, A  | 232 | 6 | 3.3 | 96 |                      |                    |

|     |   |     |     |    |                      |                    |     |   |     |     |    |                      |                    |
|-----|---|-----|-----|----|----------------------|--------------------|-----|---|-----|-----|----|----------------------|--------------------|
| 233 | 6 | 3.3 | 97  | 15 | US-10-424-599-254607 | Sequence 254607, A | 306 | 6 | 3.3 | 127 | 15 | US-10-424-599-194345 | Sequence 194345, A |
| 234 | 6 | 3.3 | 97  | 16 | US-10-437-963-171922 | Sequence 171922, A | 307 | 6 | 3.3 | 127 | 16 | US-10-437-963-107319 | Sequence 107319, A |
| 235 | 6 | 3.3 | 97  | 16 | US-10-767-701-59874  | Sequence 59874, A  | 308 | 6 | 3.3 | 127 | 16 | US-10-437-963-167924 | Sequence 167924, A |
| 236 | 6 | 3.3 | 98  | 15 | US-10-424-599-156880 | Sequence 156880, A | 309 | 6 | 3.3 | 127 | 16 | US-10-437-963-174804 | Sequence 174804, A |
| 237 | 6 | 3.3 | 98  | 16 | US-10-437-963-126643 | Sequence 126643, A | 310 | 6 | 3.3 | 127 | 16 | US-10-437-963-195477 | Sequence 195477, A |
| 238 | 6 | 3.3 | 98  | 16 | US-10-437-963-135702 | Sequence 135702, A | 311 | 6 | 3.3 | 127 | 16 | US-10-767-701-46768  | Sequence 46768, A  |
| 239 | 6 | 3.3 | 99  | 15 | US-10-425-114-46090  | Sequence 46090, A  | 312 | 6 | 3.3 | 127 | 16 | US-10-767-701-46889  | Sequence 46889, A  |
| 240 | 6 | 3.3 | 99  | 16 | US-10-437-963-147823 | Sequence 147823, A | 313 | 6 | 3.3 | 127 | 16 | US-10-767-701-46928  | Sequence 46928, A  |
| 241 | 6 | 3.3 | 99  | 16 | US-10-437-963-202123 | Sequence 202123, A | 314 | 6 | 3.3 | 128 | 15 | US-10-424-599-222218 | Sequence 222218, A |
| 242 | 6 | 3.3 | 100 | 16 | US-10-767-701-43987  | Sequence 43987, A  | 315 | 6 | 3.3 | 128 | 16 | US-10-437-963-171590 | Sequence 171590, A |
| 243 | 6 | 3.3 | 101 | 15 | US-10-424-599-224479 | Sequence 224479, A | 316 | 6 | 3.3 | 129 | 16 | US-10-437-963-136967 | Sequence 136967, A |
| 244 | 6 | 3.3 | 101 | 15 | US-10-424-599-234345 | Sequence 234345, A | 317 | 6 | 3.3 | 129 | 16 | US-10-437-963-170530 | Sequence 170530, A |
| 245 | 6 | 3.3 | 101 | 16 | US-10-437-963-112864 | Sequence 112864, A | 318 | 6 | 3.3 | 129 | 16 | US-10-437-963-194773 | Sequence 194773, A |
| 246 | 6 | 3.3 | 101 | 16 | US-10-437-963-132375 | Sequence 132375, A | 319 | 6 | 3.3 | 129 | 16 | US-10-767-701-60943  | Sequence 60943, A  |
| 247 | 6 | 3.3 | 101 | 16 | US-10-437-963-144006 | Sequence 144006, A | 320 | 6 | 3.3 | 130 | 9  | US-09-864-761-47619  | Sequence 47619, A  |
| 248 | 6 | 3.3 | 102 | 15 | US-10-424-599-179514 | Sequence 179514, A | 321 | 6 | 3.3 | 130 | 16 | US-10-437-963-152160 | Sequence 152160, A |
| 249 | 6 | 3.3 | 102 | 16 | US-10-437-963-164135 | Sequence 164135, A | 322 | 6 | 3.3 | 130 | 16 | US-10-437-963-182481 | Sequence 182481, A |
| 250 | 6 | 3.3 | 103 | 15 | US-10-424-599-148655 | Sequence 148655, A | 323 | 6 | 3.3 | 131 | 14 | US-10-106-698-7608   | Sequence 7608, Ap  |
| 251 | 6 | 3.3 | 103 | 16 | US-10-437-963-102588 | Sequence 102588, A | 324 | 6 | 3.3 | 131 | 15 | US-10-424-599-197396 | Sequence 197396, A |
| 252 | 6 | 3.3 | 103 | 16 | US-10-437-963-174574 | Sequence 174574, A | 325 | 6 | 3.3 | 131 | 16 | US-10-437-963-200756 | Sequence 200756, A |
| 253 | 6 | 3.3 | 104 | 15 | US-10-767-701-44551  | Sequence 44551, A  | 326 | 6 | 3.3 | 132 | 15 | US-10-424-599-157787 | Sequence 157787, A |
| 254 | 6 | 3.3 | 104 | 16 | US-10-424-599-175275 | Sequence 175275, A | 327 | 6 | 3.3 | 132 | 15 | US-10-424-599-237093 | Sequence 237093, A |
| 255 | 6 | 3.3 | 104 | 16 | US-10-767-701-36858  | Sequence 36858, A  | 328 | 6 | 3.3 | 133 | 16 | US-10-767-701-60171  | Sequence 60171, A  |
| 256 | 6 | 3.3 | 105 | 13 | US-10-135-322-6      | Sequence 6, Appli  | 329 | 6 | 3.3 | 134 | 9  | US-09-802-077-3      | Sequence 3, Appli  |
| 257 | 6 | 3.3 | 105 | 16 | US-10-437-963-153127 | Sequence 153127, A | 330 | 6 | 3.3 | 134 | 9  | US-09-802-096-3      | Sequence 3, Appli  |
| 258 | 6 | 3.3 | 106 | 15 | US-10-437-963-154563 | Sequence 154563, A | 331 | 6 | 3.3 | 134 | 9  | US-09-186-276B-48    | Sequence 48, Appli |
| 259 | 6 | 3.3 | 108 | 15 | US-10-424-599-246300 | Sequence 246300, A | 332 | 6 | 3.3 | 134 | 9  | US-09-186-188B-48    | Sequence 48, Appli |
| 260 | 6 | 3.3 | 108 | 15 | US-10-425-114-48223  | Sequence 48223, A  | 333 | 6 | 3.3 | 134 | 10 | US-09-925-179-3      | Sequence 3, Appli  |
| 261 | 6 | 3.3 | 108 | 15 | US-10-311-129-29     | Sequence 29, Appli | 334 | 6 | 3.3 | 134 | 14 | US-10-253-007-48     | Sequence 48, Appli |
| 262 | 6 | 3.3 | 108 | 16 | US-10-437-963-114886 | Sequence 114886, A | 335 | 6 | 3.3 | 136 | 16 | US-10-437-963-180875 | Sequence 180875, A |
| 263 | 6 | 3.3 | 108 | 16 | US-10-437-963-144765 | Sequence 144765, A | 336 | 6 | 3.3 | 136 | 16 | US-10-437-963-203306 | Sequence 203306, A |
| 264 | 6 | 3.3 | 108 | 16 | US-10-437-963-148967 | Sequence 148967, A | 337 | 6 | 3.3 | 137 | 16 | US-10-437-963-203306 | Sequence 203306, A |
| 265 | 6 | 3.3 | 109 | 15 | US-10-424-599-221195 | Sequence 221195, A | 338 | 6 | 3.3 | 137 | 15 | US-10-424-599-154925 | Sequence 154925, A |
| 266 | 6 | 3.3 | 109 | 16 | US-10-767-701-36150  | Sequence 36150, A  | 339 | 6 | 3.3 | 137 | 15 | US-10-424-599-283689 | Sequence 283689, A |
| 267 | 6 | 3.3 | 110 | 15 | US-10-251-085B-173   | Sequence 173, App  | 340 | 6 | 3.3 | 137 | 16 | US-10-767-701-34943  | Sequence 34943, A  |
| 268 | 6 | 3.3 | 110 | 16 | US-10-437-963-180146 | Sequence 180146, A | 341 | 6 | 3.3 | 138 | 16 | US-10-437-963-167524 | Sequence 167524, A |
| 269 | 6 | 3.3 | 110 | 16 | US-10-737-252-173    | Sequence 173, App  | 342 | 6 | 3.3 | 140 | 15 | US-10-424-599-280069 | Sequence 280069, A |
| 270 | 6 | 3.3 | 110 | 16 | US-10-393-134-1      | Sequence 1, Appli  | 343 | 6 | 3.3 | 141 | 15 | US-10-424-599-214807 | Sequence 214807, A |
| 271 | 6 | 3.3 | 111 | 16 | US-10-437-963-104669 | Sequence 104669, A | 344 | 6 | 3.3 | 142 | 16 | US-10-437-963-147940 | Sequence 147940, A |
| 272 | 6 | 3.3 | 112 | 9  | US-09-764-869-1081   | Sequence 1081, Ap  | 345 | 6 | 3.3 | 143 | 16 | US-10-767-701-32972  | Sequence 32972, A  |
| 273 | 6 | 3.3 | 112 | 14 | US-10-091-504-1081   | Sequence 1081, Ap  | 346 | 6 | 3.3 | 144 | 15 | US-10-424-599-163918 | Sequence 163918, A |
| 274 | 6 | 3.3 | 112 | 15 | US-10-227-577-1081   | Sequence 1081, Ap  | 347 | 6 | 3.3 | 144 | 16 | US-10-437-963-108724 | Sequence 108724, A |
| 275 | 6 | 3.3 | 112 | 15 | US-10-251-085B-169   | Sequence 169, App  | 348 | 6 | 3.3 | 145 | 15 | US-10-460-594-55     | Sequence 55, Appli |
| 276 | 6 | 3.3 | 112 | 16 | US-10-737-252-169    | Sequence 169, App  | 349 | 6 | 3.3 | 145 | 15 | US-10-276-774-1926   | Sequence 1926, Ap  |
| 277 | 6 | 3.3 | 113 | 14 | US-10-128-714-8550   | Sequence 8550, App | 349 | 6 | 3.3 | 146 | 15 | US-10-158-057-183    | Sequence 183, App  |
| 278 | 6 | 3.3 | 113 | 14 | US-10-094-749-2529   | Sequence 8550, Ap  | 350 | 6 | 3.3 | 146 | 15 | US-10-282-122A-63974 | Sequence 63974, A  |
| 279 | 6 | 3.3 | 113 | 14 | US-10-108-260A-2856  | Sequence 2529, Ap  | 351 | 6 | 3.3 | 146 | 16 | US-10-437-963-159677 | Sequence 159677, A |
| 280 | 6 | 3.3 | 113 | 15 | US-10-108-260A-2856  | Sequence 2856, Ap  | 352 | 6 | 3.3 | 147 | 15 | US-10-282-122A-61864 | Sequence 61864, A  |
| 281 | 6 | 3.3 | 113 | 16 | US-10-437-963-156315 | Sequence 156315, A | 353 | 6 | 3.3 | 147 | 16 | US-10-437-963-203492 | Sequence 203492, A |
| 282 | 6 | 3.3 | 113 | 16 | US-10-437-963-162303 | Sequence 162303, A | 354 | 6 | 3.3 | 147 | 16 | US-10-767-701-52198  | Sequence 52198, A  |
| 283 | 6 | 3.3 | 114 | 15 | US-10-424-599-275435 | Sequence 275435, A | 355 | 6 | 3.3 | 148 | 16 | US-10-424-599-204303 | Sequence 204303, A |
| 284 | 6 | 3.3 | 114 | 16 | US-10-437-963-124875 | Sequence 124875, A | 356 | 6 | 3.3 | 148 | 16 | US-10-424-599-204303 | Sequence 204303, A |
| 285 | 6 | 3.3 | 116 | 15 | US-10-437-963-122955 | Sequence 122955, A | 357 | 6 | 3.3 | 148 | 16 | US-10-767-701-61789  | Sequence 61789, A  |
| 286 | 6 | 3.3 | 118 | 15 | US-10-424-599-262788 | Sequence 262788, A | 358 | 6 | 3.3 | 149 | 14 | US-10-233-926-4      | Sequence 4, Appli  |
| 287 | 6 | 3.3 | 118 | 16 | US-10-437-963-122480 | Sequence 122480, A | 359 | 6 | 3.3 | 149 | 14 | US-10-369-493-18745  | Sequence 18745, A  |
| 288 | 6 | 3.3 | 118 | 16 | US-10-437-963-139120 | Sequence 139120, A | 360 | 6 | 3.3 | 149 | 15 | US-10-424-599-162303 | Sequence 162303, A |
| 289 | 6 | 3.3 | 118 | 16 | US-10-767-701-32567  | Sequence 32567, A  | 361 | 6 | 3.3 | 149 | 16 | US-10-437-963-136643 | Sequence 136643, A |
| 290 | 6 | 3.3 | 119 | 16 | US-10-437-963-109271 | Sequence 109271, A | 362 | 6 | 3.3 | 150 | 9  | US-09-883-985-11     | Sequence 11, Appli |
| 291 | 6 | 3.3 | 120 | 15 | US-10-424-599-146939 | Sequence 146939, A | 363 | 6 | 3.3 | 150 | 14 | US-10-290-142-5      | Sequence 5, Appli  |
| 292 | 6 | 3.3 | 120 | 15 | US-10-424-599-232985 | Sequence 232985, A | 364 | 6 | 3.3 | 150 | 14 | US-10-290-143-4      | Sequence 4, Appli  |
| 293 | 6 | 3.3 | 120 | 15 | US-10-424-599-277939 | Sequence 277939, A | 365 | 6 | 3.3 | 150 | 16 | US-10-437-963-182530 | Sequence 182530, A |
| 294 | 6 | 3.3 | 121 | 9  | US-09-864-761-34580  | Sequence 34580, A  | 366 | 6 | 3.3 | 151 | 15 | US-10-425-114-56523  | Sequence 56523, A  |
| 295 | 6 | 3.3 | 121 | 9  | US-09-920-171-2      | Sequence 2, Appli  | 367 | 6 | 3.3 | 151 | 16 | US-10-437-963-119514 | Sequence 119514, A |
| 296 | 6 | 3.3 | 121 | 14 | US-10-113-996-2      | Sequence 2, Appli  | 368 | 6 | 3.3 | 151 | 16 | US-10-437-963-121948 | Sequence 121948, A |
| 297 | 6 | 3.3 | 122 | 15 | US-10-424-599-216013 | Sequence 216013, A | 369 | 6 | 3.3 | 151 | 16 | US-10-437-963-177780 | Sequence 177780, A |
| 298 | 6 | 3.3 | 122 | 16 | US-10-767-701-41124  | Sequence 41124, A  | 370 | 6 | 3.3 | 152 | 14 | US-10-205-194-57     | Sequence 57, Appli |
| 299 | 6 | 3.3 | 125 | 15 | US-10-108-260A-2869  | Sequence 2869, Ap  | 371 | 6 | 3.3 | 152 | 15 | US-10-424-599-209356 | Sequence 209356, A |
| 300 | 6 | 3.3 | 125 | 15 | US-10-424-599-253721 | Sequence 253721, A | 372 | 6 | 3.3 | 152 | 15 | US-10-424-599-210951 | Sequence 210951, A |
| 301 | 6 | 3.3 | 125 | 16 | US-10-437-963-135370 | Sequence 135370, A | 373 | 6 | 3.3 | 152 | 16 | US-10-437-963-151821 | Sequence 151821, A |
| 302 | 6 | 3.3 | 126 | 15 | US-10-424-599-155585 | Sequence 155585, A | 374 | 6 | 3.3 | 152 | 16 | US-10-437-963-163001 | Sequence 163001, A |
| 303 | 6 | 3.3 | 126 | 15 | US-10-425-114-56414  | Sequence 56414, A  | 375 | 6 | 3.3 | 152 | 16 | US-10-767-701-32071  | Sequence 32071, A  |
| 304 | 6 | 3.3 | 126 | 16 | US-10-437-963-112575 | Sequence 112575, A | 376 | 6 | 3.3 | 153 | 15 | US-10-424-599-159963 | Sequence 159963, A |
| 305 | 6 | 3.3 | 126 | 16 | US-10-437-963-138517 | Sequence 138517, A | 377 | 6 | 3.3 | 153 | 15 | US-10-424-599-163211 | Sequence 163211, A |
|     |   |     |     |    |                      |                    | 378 | 6 | 3.3 | 153 | 15 | US-10-424-599-248721 | Sequence 248721, A |

|     |   |     |     |    |                      |                    |     |   |     |     |    |                      |                    |
|-----|---|-----|-----|----|----------------------|--------------------|-----|---|-----|-----|----|----------------------|--------------------|
| 379 | 6 | 3.3 | 153 | 15 | US-10-425-114-67718  | Sequence 67718, A  | 452 | 6 | 3.3 | 180 | 16 | US-10-767-701-39277  | Sequence 39277, A  |
| 380 | 6 | 3.3 | 153 | 15 | US-10-399-241A-14    | Sequence 14, Appl  | 453 | 6 | 3.3 | 180 | 16 | US-10-767-701-57218  | Sequence 57218, A  |
| 381 | 6 | 3.3 | 154 | 15 | US-10-424-599-170469 | Sequence 170469, A | 454 | 6 | 3.3 | 181 | 14 | US-10-017-161-1378   | Sequence 1378, Ap  |
| 382 | 6 | 3.3 | 154 | 15 | US-10-424-599-197868 | Sequence 197868, A | 455 | 6 | 3.3 | 181 | 14 | US-10-292-798-1128   | Sequence 1128, Ap  |
| 383 | 6 | 3.3 | 154 | 16 | US-10-437-963-117935 | Sequence 117935, A | 456 | 6 | 3.3 | 181 | 15 | US-10-425-114-57659  | Sequence 57659, A  |
| 384 | 6 | 3.3 | 154 | 16 | US-10-437-963-197938 | Sequence 197938, A | 457 | 6 | 3.3 | 181 | 16 | US-10-767-701-59959  | Sequence 59959, A  |
| 385 | 6 | 3.3 | 154 | 16 | US-10-437-963-197938 | Sequence 197938, A | 458 | 6 | 3.3 | 182 | 16 | US-10-437-963-124285 | Sequence 124285, A |
| 386 | 6 | 3.3 | 155 | 15 | US-10-424-599-219149 | Sequence 51949, A  | 459 | 6 | 3.3 | 183 | 15 | US-10-424-599-240010 | Sequence 240010, A |
| 387 | 6 | 3.3 | 155 | 15 | US-10-424-599-213389 | Sequence 213389, A | 460 | 6 | 3.3 | 183 | 15 | US-10-437-963-174617 | Sequence 174617, A |
| 388 | 6 | 3.3 | 155 | 16 | US-10-437-963-110237 | Sequence 110237, A | 461 | 6 | 3.3 | 183 | 16 | US-10-767-701-33267  | Sequence 33267, A  |
| 389 | 6 | 3.3 | 155 | 16 | US-10-437-963-174153 | Sequence 174153, A | 462 | 6 | 3.3 | 183 | 16 | US-10-767-701-45556  | Sequence 45556, A  |
| 390 | 6 | 3.3 | 156 | 15 | US-10-424-599-188143 | Sequence 188143, A | 463 | 6 | 3.3 | 183 | 16 | US-10-424-599-220518 | Sequence 220518, A |
| 391 | 6 | 3.3 | 156 | 15 | US-10-424-599-266125 | Sequence 266125, A | 464 | 6 | 3.3 | 184 | 15 | US-10-092-900A-238   | Sequence 238, App  |
| 392 | 6 | 3.3 | 156 | 16 | US-10-767-701-34285  | Sequence 34285, A  | 465 | 6 | 3.3 | 184 | 15 | US-10-437-963-145269 | Sequence 145269, A |
| 393 | 6 | 3.3 | 157 | 16 | US-10-437-963-149833 | Sequence 149833, A | 466 | 6 | 3.3 | 185 | 16 | US-10-093-455-312    | Sequence 312, App  |
| 394 | 6 | 3.3 | 157 | 16 | US-10-437-963-153042 | Sequence 153042, A | 467 | 6 | 3.3 | 185 | 14 | US-10-295-403-38     | Sequence 38, Appl  |
| 395 | 6 | 3.3 | 158 | 14 | US-10-369-493-19755  | Sequence 19755, A  | 468 | 6 | 3.3 | 185 | 14 | US-10-225-066A-618   | Sequence 618, App  |
| 396 | 6 | 3.3 | 158 | 15 | US-10-425-114-46108  | Sequence 46108, A  | 469 | 6 | 3.3 | 185 | 15 | US-10-374-780A-2504  | Sequence 2504, Ap  |
| 397 | 6 | 3.3 | 158 | 16 | US-10-437-963-109346 | Sequence 109346, A | 470 | 6 | 3.3 | 185 | 15 | US-10-412-699B-674   | Sequence 674, App  |
| 398 | 6 | 3.3 | 158 | 16 | US-10-437-963-154676 | Sequence 154676, A | 471 | 6 | 3.3 | 186 | 15 | US-10-425-114-62639  | Sequence 62639, A  |
| 399 | 6 | 3.3 | 158 | 16 | US-10-437-963-171703 | Sequence 171703, A | 472 | 6 | 3.3 | 186 | 15 | US-10-389-566-823    | Sequence 823, App  |
| 400 | 6 | 3.3 | 159 | 9  | US-09-746-284-15     | Sequence 16, Appl  | 473 | 6 | 3.3 | 188 | 15 | US-10-425-114-67096  | Sequence 67096, A  |
| 401 | 6 | 3.3 | 159 | 15 | US-10-424-599-210722 | Sequence 210722, A | 474 | 6 | 3.3 | 188 | 16 | US-10-437-963-198685 | Sequence 198685, A |
| 402 | 6 | 3.3 | 159 | 16 | US-10-437-963-115997 | Sequence 115997, A | 475 | 6 | 3.3 | 190 | 9  | US-09-867-550-668    | Sequence 668, App  |
| 403 | 6 | 3.3 | 160 | 15 | US-10-424-599-143062 | Sequence 143062, A | 476 | 6 | 3.3 | 190 | 16 | US-10-437-963-124435 | Sequence 124435, A |
| 404 | 6 | 3.3 | 160 | 16 | US-10-767-701-34756  | Sequence 34756, A  | 477 | 6 | 3.3 | 190 | 16 | US-10-437-963-132486 | Sequence 132486, A |
| 405 | 6 | 3.3 | 161 | 15 | US-10-282-122A-62658 | Sequence 62658, A  | 478 | 6 | 3.3 | 191 | 15 | US-10-424-599-160161 | Sequence 160161, A |
| 406 | 6 | 3.3 | 161 | 15 | US-10-282-122A-64408 | Sequence 64408, A  | 479 | 6 | 3.3 | 191 | 15 | US-10-424-599-178644 | Sequence 178644, A |
| 407 | 6 | 3.3 | 161 | 15 | US-10-424-599-243761 | Sequence 243761, A | 480 | 6 | 3.3 | 192 | 16 | US-10-767-701-40683  | Sequence 40683, A  |
| 408 | 6 | 3.3 | 161 | 16 | US-10-437-963-108823 | Sequence 108823, A | 481 | 6 | 3.3 | 193 | 9  | US-09-925-298-706    | Sequence 706, App  |
| 409 | 6 | 3.3 | 162 | 15 | US-10-767-701-61708  | Sequence 61708, A  | 482 | 6 | 3.3 | 193 | 14 | US-10-102-806-706    | Sequence 706, App  |
| 410 | 6 | 3.3 | 162 | 16 | US-10-425-114-41899  | Sequence 41899, A  | 483 | 6 | 3.3 | 193 | 14 | US-10-106-698-4763   | Sequence 4763, Ap  |
| 411 | 6 | 3.3 | 163 | 16 | US-10-437-963-125479 | Sequence 125479, A | 484 | 6 | 3.3 | 193 | 16 | US-10-437-963-163056 | Sequence 163056, A |
| 412 | 6 | 3.3 | 163 | 16 | US-10-437-963-168736 | Sequence 168736, A | 485 | 6 | 3.3 | 194 | 15 | US-10-425-114-38145  | Sequence 38145, A  |
| 413 | 6 | 3.3 | 163 | 16 | US-10-437-963-200787 | Sequence 200787, A | 486 | 6 | 3.3 | 194 | 15 | US-10-236-115-1364   | Sequence 1364, Ap  |
| 414 | 6 | 3.3 | 163 | 16 | US-10-767-701-33889  | Sequence 33889, A  | 487 | 6 | 3.3 | 194 | 16 | US-10-437-963-177828 | Sequence 177828, A |
| 415 | 6 | 3.3 | 164 | 14 | US-10-369-493-19892  | Sequence 19892, A  | 488 | 6 | 3.3 | 194 | 16 | US-10-437-963-177832 | Sequence 177832, A |
| 416 | 6 | 3.3 | 164 | 16 | US-10-437-963-151969 | Sequence 151969, A | 489 | 6 | 3.3 | 194 | 16 | US-10-437-963-185204 | Sequence 185204, A |
| 417 | 6 | 3.3 | 165 | 16 | US-10-437-963-185203 | Sequence 185203, A | 490 | 6 | 3.3 | 194 | 16 | US-10-437-963-192508 | Sequence 192508, A |
| 418 | 6 | 3.3 | 166 | 14 | US-10-080-170-278    | Sequence 278, App  | 491 | 6 | 3.3 | 195 | 9  | US-09-864-761-36421  | Sequence 36421, A  |
| 419 | 6 | 3.3 | 166 | 15 | US-10-424-599-245345 | Sequence 245345, A | 492 | 6 | 3.3 | 195 | 15 | US-10-108-260A-4832  | Sequence 4832, Ap  |
| 420 | 6 | 3.3 | 166 | 16 | US-10-080-170-278    | Sequence 278, App  | 493 | 6 | 3.3 | 196 | 9  | US-09-886-468-15     | Sequence 15, Appl  |
| 421 | 6 | 3.3 | 166 | 17 | US-10-468-356-278    | Sequence 278, App  | 494 | 6 | 3.3 | 196 | 9  | US-09-889-468-15     | Sequence 15, Appl  |
| 422 | 6 | 3.3 | 167 | 11 | US-09-864-408A-2174  | Sequence 2174, App | 495 | 6 | 3.3 | 196 | 15 | US-10-425-114-44993  | Sequence 44993, A  |
| 423 | 6 | 3.3 | 167 | 15 | US-10-282-122A-54352 | Sequence 54352, A  | 496 | 6 | 3.3 | 197 | 9  | US-09-994-185-4      | Sequence 2, Appl   |
| 424 | 6 | 3.3 | 167 | 15 | US-10-425-114-67362  | Sequence 67362, A  | 497 | 6 | 3.3 | 197 | 14 | US-10-131-686A-2     | Sequence 2, Appl   |
| 425 | 6 | 3.3 | 167 | 16 | US-10-437-963-108870 | Sequence 108870, A | 498 | 6 | 3.3 | 197 | 14 | US-10-131-686A-45    | Sequence 45, Appl  |
| 426 | 6 | 3.3 | 168 | 15 | US-10-425-114-62191  | Sequence 62191, A  | 499 | 6 | 3.3 | 197 | 14 | US-10-131-686A-51    | Sequence 51, Appl  |
| 427 | 6 | 3.3 | 169 | 14 | US-10-104-047-9909   | Sequence 3909, App | 500 | 6 | 3.3 | 197 | 14 | US-10-032-585-7743   | Sequence 7743, Ap  |
| 428 | 6 | 3.3 | 169 | 15 | US-10-363-616-361    | Sequence 361, App  | 501 | 6 | 3.3 | 197 | 15 | US-10-282-122A-42890 | Sequence 42890, A  |
| 429 | 6 | 3.3 | 169 | 16 | US-10-767-701-33653  | Sequence 33653, A  | 502 | 6 | 3.3 | 197 | 15 | US-10-282-122A-56284 | Sequence 56284, A  |
| 430 | 6 | 3.3 | 170 | 14 | US-10-012-819-24     | Sequence 24, Appl  | 503 | 6 | 3.3 | 197 | 15 | US-10-425-114-53690  | Sequence 53690, A  |
| 431 | 6 | 3.3 | 170 | 16 | US-10-437-963-123554 | Sequence 123554, A | 504 | 6 | 3.3 | 197 | 16 | US-10-437-963-112713 | Sequence 112713, A |
| 432 | 6 | 3.3 | 170 | 16 | US-10-437-963-155742 | Sequence 155742, A | 505 | 6 | 3.3 | 198 | 14 | US-10-131-686A-52    | Sequence 52, Appl  |
| 433 | 6 | 3.3 | 171 | 16 | US-10-767-701-39779  | Sequence 39779, A  | 506 | 6 | 3.3 | 198 | 14 | US-10-131-686A-53    | Sequence 53, Appl  |
| 434 | 6 | 3.3 | 171 | 15 | US-10-108-260A-4136  | Sequence 4136, Ap  | 507 | 6 | 3.3 | 198 | 16 | US-10-437-963-184042 | Sequence 184042, A |
| 435 | 6 | 3.3 | 171 | 15 | US-10-424-599-155473 | Sequence 155473, A | 508 | 6 | 3.3 | 199 | 14 | US-10-131-686A-46    | Sequence 46, Appl  |
| 436 | 6 | 3.3 | 172 | 16 | US-10-437-963-177319 | Sequence 177319, A | 509 | 6 | 3.3 | 199 | 14 | US-10-131-686A-48    | Sequence 48, Appl  |
| 437 | 6 | 3.3 | 173 | 16 | US-10-437-963-127931 | Sequence 127931, A | 510 | 6 | 3.3 | 199 | 15 | US-10-289-762-595    | Sequence 595, App  |
| 438 | 6 | 3.3 | 173 | 16 | US-10-437-963-127931 | Sequence 127931, A | 511 | 6 | 3.3 | 201 | 11 | US-10-289-762-595    | Sequence 505, App  |
| 439 | 6 | 3.3 | 175 | 16 | US-10-437-963-137844 | Sequence 137844, A | 512 | 6 | 3.3 | 201 | 15 | US-10-424-599-273773 | Sequence 273773, A |
| 440 | 6 | 3.3 | 175 | 16 | US-10-437-963-179308 | Sequence 179308, A | 513 | 6 | 3.3 | 201 | 16 | US-10-437-963-166716 | Sequence 166716, A |
| 441 | 6 | 3.3 | 175 | 16 | US-10-437-963-191373 | Sequence 191373, A | 514 | 6 | 3.3 | 202 | 15 | US-10-425-114-53711  | Sequence 53711, A  |
| 442 | 6 | 3.3 | 176 | 16 | US-10-437-963-104324 | Sequence 104324, A | 515 | 6 | 3.3 | 202 | 16 | US-10-767-701-57454  | Sequence 57454, A  |
| 443 | 6 | 3.3 | 176 | 16 | US-10-437-963-105030 | Sequence 105030, A | 516 | 6 | 3.3 | 203 | 16 | US-10-437-963-157433 | Sequence 157433, A |
| 444 | 6 | 3.3 | 176 | 16 | US-10-767-701-58229  | Sequence 58229, A  | 517 | 6 | 3.3 | 204 | 10 | US-09-892-877-387    | Sequence 387, App  |
| 445 | 6 | 3.3 | 177 | 16 | US-10-437-963-109348 | Sequence 109348, A | 518 | 6 | 3.3 | 204 | 10 | US-09-948-783-387    | Sequence 387, App  |
| 446 | 6 | 3.3 | 177 | 16 | US-10-437-963-191610 | Sequence 191610, A | 519 | 6 | 3.3 | 204 | 15 | US-10-424-599-172260 | Sequence 172260, A |
| 447 | 6 | 3.3 | 178 | 15 | US-10-425-114-48081  | Sequence 48081, A  | 520 | 6 | 3.3 | 204 | 15 | US-10-424-599-202746 | Sequence 202746, A |
| 448 | 6 | 3.3 | 179 | 15 | US-10-425-114-62328  | Sequence 62328, A  | 521 | 6 | 3.3 | 204 | 15 | US-10-425-114-43377  | Sequence 43377, A  |
| 449 | 6 | 3.3 | 180 | 11 | US-09-833-245-1716   | Sequence 1716, Ap  | 522 | 6 | 3.3 | 204 | 15 | US-10-425-114-47766  | Sequence 47766, A  |
| 450 | 6 | 3.3 | 180 | 11 | US-09-833-245-1718   | Sequence 1718, Ap  | 523 | 6 | 3.3 | 204 | 16 | US-10-437-963-179624 | Sequence 179624, A |
| 451 | 6 | 3.3 | 180 | 14 | US-10-174-693-313    | Sequence 313, App  | 524 | 6 | 3.3 | 204 | 16 | US-10-767-701-54752  | Sequence 54752, A  |

|     |   |     |     |    |                       |                    |     |   |     |     |    |                      |                    |
|-----|---|-----|-----|----|-----------------------|--------------------|-----|---|-----|-----|----|----------------------|--------------------|
| 525 | 6 | 3.3 | 207 | 16 | US-10-437-963-1145444 | Sequence 145444,   | 598 | 6 | 3.3 | 236 | 15 | US-10-424-599-282294 | Sequence 282294,   |
| 526 | 6 | 3.3 | 208 | 14 | US-10-096-327-2       | Sequence 2, Appli  | 599 | 6 | 3.3 | 237 | 15 | US-10-282-122A-64317 | Sequence 64317, A  |
| 527 | 6 | 3.3 | 208 | 14 | US-10-138-158-18      | Sequence 18, Appl  | 600 | 6 | 3.3 | 237 | 15 | US-10-282-122A-66354 | Sequence 66354, A  |
| 528 | 6 | 3.3 | 208 | 14 | US-10-189-360-5       | Sequence 5, Appli  | 601 | 6 | 3.3 | 237 | 15 | US-10-424-599-236491 | Sequence 236491, A |
| 529 | 6 | 3.3 | 208 | 14 | US-10-395-540-2       | Sequence 2, Appli  | 602 | 6 | 3.3 | 237 | 15 | US-10-425-114-57223  | Sequence 57223, A  |
| 530 | 6 | 3.3 | 208 | 15 | US-10-428-540-38      | Sequence 38, Appli | 603 | 6 | 3.3 | 238 | 15 | US-10-282-122A-47691 | Sequence 47691, A  |
| 531 | 6 | 3.3 | 208 | 15 | US-10-282-122A-48748  | Sequence 48748, A  | 604 | 6 | 3.3 | 238 | 15 | US-10-282-122A-50015 | Sequence 50015, A  |
| 532 | 6 | 3.3 | 208 | 15 | US-10-425-114-48916   | Sequence 48916, A  | 605 | 6 | 3.3 | 238 | 15 | US-10-282-122A-52326 | Sequence 52326, A  |
| 533 | 6 | 3.3 | 208 | 16 | US-10-437-963-160553  | Sequence 160553,   | 606 | 6 | 3.3 | 238 | 15 | US-10-424-599-217919 | Sequence 217919, A |
| 534 | 6 | 3.3 | 208 | 16 | US-10-437-963-198586  | Sequence 198586,   | 607 | 6 | 3.3 | 239 | 14 | US-10-156-761-7951   | Sequence 7951, Ap  |
| 535 | 6 | 3.3 | 209 | 16 | US-10-437-963-190342  | Sequence 190342,   | 608 | 6 | 3.3 | 240 | 14 | US-10-143-897-2      | Sequence 2, Appli  |
| 536 | 6 | 3.3 | 209 | 16 | US-10-437-963-199319  | Sequence 199319,   | 609 | 6 | 3.3 | 240 | 14 | US-10-143-897-12     | Sequence 12, Appl  |
| 537 | 6 | 3.3 | 210 | 15 | US-10-282-122A-53481  | Sequence 53481, A  | 610 | 6 | 3.3 | 240 | 15 | US-10-424-599-236490 | Sequence 236490,   |
| 538 | 6 | 3.3 | 210 | 15 | US-10-424-599-246223  | Sequence 246223,   | 611 | 6 | 3.3 | 240 | 15 | US-10-425-114-57701  | Sequence 57701, A  |
| 539 | 6 | 3.3 | 210 | 16 | US-10-437-963-172516  | Sequence 172516,   | 612 | 6 | 3.3 | 241 | 16 | US-10-437-963-137063 | Sequence 137063,   |
| 540 | 6 | 3.3 | 211 | 14 | US-10-369-493-14049   | Sequence 14049, A  | 613 | 6 | 3.3 | 242 | 10 | US-09-880-748-1830   | Sequence 1830, Ap  |
| 541 | 6 | 3.3 | 211 | 15 | US-10-425-114-54325   | Sequence 54325, A  | 614 | 6 | 3.3 | 242 | 14 | US-10-029-386-32982  | Sequence 32982, A  |
| 542 | 6 | 3.3 | 211 | 16 | US-10-437-963-118968  | Sequence 118968,   | 615 | 6 | 3.3 | 242 | 14 | US-10-293-418-1830   | Sequence 1830, Ap  |
| 543 | 6 | 3.3 | 212 | 15 | US-10-425-114-70708   | Sequence 70708, A  | 616 | 6 | 3.3 | 242 | 16 | US-10-437-963-203736 | Sequence 203736,   |
| 544 | 6 | 3.3 | 213 | 16 | US-10-767-701-32228   | Sequence 32228, A  | 617 | 6 | 3.3 | 243 | 15 | US-10-369-435-4      | Sequence 4, Appli  |
| 545 | 6 | 3.3 | 214 | 16 | US-10-437-963-111588  | Sequence 111588,   | 618 | 6 | 3.3 | 245 | 16 | US-10-437-963-126783 | Sequence 126783,   |
| 546 | 6 | 3.3 | 215 | 14 | US-10-094-749-1759    | Sequence 1759, Ap  | 619 | 6 | 3.3 | 245 | 16 | US-10-437-963-150619 | Sequence 150619,   |
| 547 | 6 | 3.3 | 215 | 15 | US-10-425-114-50866   | Sequence 50866, A  | 620 | 6 | 3.3 | 245 | 16 | US-10-437-963-159190 | Sequence 159190,   |
| 548 | 6 | 3.3 | 215 | 15 | US-10-425-114-62269   | Sequence 62269, A  | 621 | 6 | 3.3 | 246 | 14 | US-10-369-493-17865  | Sequence 17865, A  |
| 549 | 6 | 3.3 | 215 | 16 | US-10-437-963-139138  | Sequence 139138,   | 622 | 6 | 3.3 | 246 | 15 | US-10-282-122A-68486 | Sequence 68486, A  |
| 550 | 6 | 3.3 | 215 | 16 | US-10-437-963-130206  | Sequence 190206,   | 623 | 6 | 3.3 | 246 | 15 | US-10-425-114-62546  | Sequence 62546, A  |
| 551 | 6 | 3.3 | 215 | 16 | US-10-437-963-131184  | Sequence 191184,   | 624 | 6 | 3.3 | 246 | 15 | US-10-276-774-2466   | Sequence 2466, Ap  |
| 552 | 6 | 3.3 | 216 | 16 | US-10-767-701-33665   | Sequence 33665, A  | 625 | 6 | 3.3 | 247 | 16 | US-10-437-963-164918 | Sequence 164918,   |
| 553 | 6 | 3.3 | 217 | 16 | US-10-437-963-177052  | Sequence 177052,   | 626 | 6 | 3.3 | 247 | 16 | US-10-437-963-166376 | Sequence 166376,   |
| 554 | 6 | 3.3 | 219 | 14 | US-10-156-761-16654   | Sequence 10654, A  | 627 | 6 | 3.3 | 247 | 16 | US-10-437-963-186474 | Sequence 186474,   |
| 555 | 6 | 3.3 | 219 | 14 | US-10-104-047-2324    | Sequence 2324, Ap  | 628 | 6 | 3.3 | 248 | 16 | US-10-437-963-193509 | Sequence 193509,   |
| 556 | 6 | 3.3 | 219 | 15 | US-10-264-049-2395    | Sequence 2395, Ap  | 629 | 6 | 3.3 | 249 | 13 | US-10-087-192-1839   | Sequence 1839, Ap  |
| 557 | 6 | 3.3 | 220 | 16 | US-10-437-963-166686  | Sequence 166686,   | 630 | 6 | 3.3 | 249 | 15 | US-10-425-114-60518  | Sequence 60518, A  |
| 558 | 6 | 3.3 | 220 | 16 | US-10-767-701-45789   | Sequence 45789, A  | 631 | 6 | 3.3 | 250 | 9  | US-09-738-626-4948   | Sequence 4948, Ap  |
| 559 | 6 | 3.3 | 221 | 9  | US-09-303-510-8       | Sequence 8, Appli  | 632 | 6 | 3.3 | 250 | 15 | US-10-282-122A-50301 | Sequence 50301, A  |
| 560 | 6 | 3.3 | 221 | 9  | US-09-303-040-8       | Sequence 8, Appli  | 633 | 6 | 3.3 | 251 | 16 | US-10-437-963-163870 | Sequence 163870,   |
| 561 | 6 | 3.3 | 221 | 15 | US-10-231-913-68      | Sequence 68, Appl  | 634 | 6 | 3.3 | 251 | 16 | US-10-767-701-43637  | Sequence 43637, A  |
| 562 | 6 | 3.3 | 221 | 15 | US-10-231-913-69      | Sequence 69, Appl  | 635 | 6 | 3.3 | 252 | 14 | US-10-145-206-115    | Sequence 115, App  |
| 563 | 6 | 3.3 | 221 | 15 | US-10-231-913-70      | Sequence 70, Appl  | 636 | 6 | 3.3 | 252 | 14 | US-10-410-647-35     | Sequence 36, Appl  |
| 564 | 6 | 3.3 | 221 | 15 | US-10-425-114-46535   | Sequence 46535, A  | 637 | 6 | 3.3 | 252 | 15 | US-10-282-762-975    | Sequence 975, App  |
| 565 | 6 | 3.3 | 221 | 15 | US-10-425-114-68801   | Sequence 68801, A  | 638 | 6 | 3.3 | 252 | 15 | US-10-425-114-48594  | Sequence 48594, A  |
| 566 | 6 | 3.3 | 221 | 16 | US-10-437-963-200734  | Sequence 200734,   | 639 | 6 | 3.3 | 253 | 14 | US-10-369-493-19538  | Sequence 19538, A  |
| 567 | 6 | 3.3 | 222 | 15 | US-10-424-599-163246  | Sequence 163246,   | 640 | 6 | 3.3 | 254 | 15 | US-10-425-114-70401  | Sequence 70401, A  |
| 568 | 6 | 3.3 | 222 | 16 | US-10-437-963-204008  | Sequence 204008,   | 641 | 6 | 3.3 | 255 | 15 | US-10-425-114-71018  | Sequence 71018, A  |
| 569 | 6 | 3.3 | 223 | 15 | US-10-425-114-65373   | Sequence 65373, A  | 642 | 6 | 3.3 | 256 | 10 | US-09-820-843A-18    | Sequence 18, Appl  |
| 570 | 6 | 3.3 | 224 | 14 | US-10-029-386-33710   | Sequence 33710, A  | 643 | 6 | 3.3 | 258 | 10 | US-09-892-877-393    | Sequence 393, App  |
| 571 | 6 | 3.3 | 224 | 15 | US-10-282-122A-55302  | Sequence 55302, A  | 644 | 6 | 3.3 | 258 | 10 | US-09-892-877-393    | Sequence 393, App  |
| 572 | 6 | 3.3 | 224 | 16 | US-10-437-963-192619  | Sequence 192619,   | 645 | 6 | 3.3 | 259 | 14 | US-10-369-493-17939  | Sequence 17939, A  |
| 573 | 6 | 3.3 | 224 | 16 | US-10-437-963-202798  | Sequence 202798,   | 646 | 6 | 3.3 | 259 | 15 | US-10-424-599-220551 | Sequence 220551,   |
| 574 | 6 | 3.3 | 225 | 15 | US-10-042-865-133     | Sequence 133, App  | 647 | 6 | 3.3 | 259 | 15 | US-10-424-599-259689 | Sequence 259689,   |
| 575 | 6 | 3.3 | 226 | 14 | US-10-029-386-32061   | Sequence 32061, A  | 648 | 6 | 3.3 | 259 | 16 | US-10-437-963-119942 | Sequence 119942,   |
| 576 | 6 | 3.3 | 226 | 14 | US-10-029-386-34006   | Sequence 34006, A  | 649 | 6 | 3.3 | 260 | 16 | US-10-437-963-133130 | Sequence 133130,   |
| 577 | 6 | 3.3 | 227 | 15 | US-10-424-599-240397  | Sequence 240397,   | 650 | 6 | 3.3 | 261 | 15 | US-10-424-599-219972 | Sequence 219972,   |
| 578 | 6 | 3.3 | 227 | 16 | US-10-437-963-114049  | Sequence 114049,   | 651 | 6 | 3.3 | 261 | 15 | US-10-425-114-72310  | Sequence 72310, A  |
| 579 | 6 | 3.3 | 228 | 16 | US-10-767-701-51602   | Sequence 51602, A  | 652 | 6 | 3.3 | 261 | 16 | US-10-437-963-179880 | Sequence 179880,   |
| 580 | 6 | 3.3 | 228 | 15 | US-10-424-599-240400  | Sequence 240400,   | 653 | 6 | 3.3 | 262 | 15 | US-10-425-114-64500  | Sequence 64500, A  |
| 581 | 6 | 3.3 | 228 | 16 | US-10-437-963-143729  | Sequence 143729,   | 654 | 6 | 3.3 | 262 | 16 | US-10-437-963-135251 | Sequence 135251,   |
| 582 | 6 | 3.3 | 229 | 15 | US-10-425-114-59306   | Sequence 59306, A  | 655 | 6 | 3.3 | 262 | 16 | US-10-767-701-42290  | Sequence 42290, A  |
| 583 | 6 | 3.3 | 229 | 16 | US-10-437-963-120982  | Sequence 120982,   | 656 | 6 | 3.3 | 263 | 9  | US-09-974-298-110    | Sequence 110, App  |
| 584 | 6 | 3.3 | 230 | 16 | US-10-437-963-134358  | Sequence 134358,   | 657 | 6 | 3.3 | 263 | 15 | US-10-425-114-67367  | Sequence 67367, A  |
| 585 | 6 | 3.3 | 230 | 16 | US-10-437-963-187846  | Sequence 187846,   | 658 | 6 | 3.3 | 263 | 16 | US-10-437-963-145396 | Sequence 145396,   |
| 586 | 6 | 3.3 | 231 | 15 | US-10-425-114-42258   | Sequence 42258, A  | 659 | 6 | 3.3 | 264 | 15 | US-10-424-599-230755 | Sequence 230755,   |
| 587 | 6 | 3.3 | 231 | 15 | US-10-080-170-615     | Sequence 615, App  | 660 | 6 | 3.3 | 264 | 16 | US-10-437-963-188302 | Sequence 188302,   |
| 588 | 6 | 3.3 | 232 | 16 | US-10-080-170-615     | Sequence 615, App  | 661 | 6 | 3.3 | 265 | 13 | US-10-104-019-23     | Sequence 23, Appl  |
| 589 | 6 | 3.3 | 232 | 16 | US-10-437-963-114316  | Sequence 114316,   | 662 | 6 | 3.3 | 265 | 14 | US-10-226-065-5      | Sequence 5, Appli  |
| 590 | 6 | 3.3 | 232 | 17 | US-10-468-356-615     | Sequence 615, App  | 663 | 6 | 3.3 | 265 | 14 | US-10-104-019-23     | Sequence 23, Appl  |
| 591 | 6 | 3.3 | 233 | 15 | US-10-369-435-36      | Sequence 36, Appl  | 664 | 6 | 3.3 | 265 | 14 | US-10-438-826-23     | Sequence 23, Appl  |
| 592 | 6 | 3.3 | 233 | 15 | US-10-312-273-167     | Sequence 167, App  | 665 | 6 | 3.3 | 265 | 16 | US-10-437-963-193584 | Sequence 193584,   |
| 593 | 6 | 3.3 | 233 | 16 | US-10-437-963-117207  | Sequence 117207,   | 666 | 6 | 3.3 | 265 | 16 | US-10-737-658-5      | Sequence 5, Appli  |
| 594 | 6 | 3.3 | 233 | 16 | US-10-676-903-2       | Sequence 2, Appli  | 667 | 6 | 3.3 | 267 | 9  | US-09-815-242-11549  | Sequence 11549, A  |
| 595 | 6 | 3.3 | 235 | 16 | US-10-437-963-192520  | Sequence 192520,   | 668 | 6 | 3.3 | 267 | 15 | US-10-335-977-8770   | Sequence 8770, Ap  |
| 596 | 6 | 3.3 | 236 | 9  | US-09-746-284-6       | Sequence 6, Appli  | 669 | 6 | 3.3 | 267 | 16 | US-10-767-701-45875  | Sequence 45875, A  |
| 597 | 6 | 3.3 | 236 | 13 | US-10-087-192-318     | Sequence 318, App  | 670 | 6 | 3.3 | 268 | 16 | US-10-474-776-413    | Sequence 413, App  |





|     |   |     |     |    |                      |                    |     |   |     |     |    |                      |                    |
|-----|---|-----|-----|----|----------------------|--------------------|-----|---|-----|-----|----|----------------------|--------------------|
| 817 | 6 | 3.3 | 328 | 16 | US-10-408-765A-96    | Sequence 96, Appl  | 890 | 6 | 3.3 | 351 | 15 | US-10-425-114-64975  | Sequence 64975, A  |
| 818 | 6 | 3.3 | 328 | 16 | US-10-648-593-232    | Sequence 232, Appl | 891 | 6 | 3.3 | 351 | 16 | US-10-437-963-155446 | Sequence 155446, A |
| 819 | 6 | 3.3 | 328 | 16 | US-10-437-963-109340 | Sequence 109340, A | 892 | 6 | 3.3 | 351 | 16 | US-10-437-963-192474 | Sequence 192474, A |
| 820 | 6 | 3.3 | 328 | 16 | US-10-706-791-9      | Sequence 9, Appli  | 893 | 6 | 3.3 | 351 | 16 | US-10-437-963-197603 | Sequence 197603, A |
| 821 | 6 | 3.3 | 328 | 16 | US-10-767-701-38418  | Sequence 38418, A  | 894 | 6 | 3.3 | 351 | 16 | US-10-748-789-2      | Sequence 2, Appli  |
| 822 | 6 | 3.3 | 329 | 14 | US-10-233-926-24     | Sequence 24, Appl  | 895 | 6 | 3.3 | 352 | 14 | US-10-094-749-2300   | Sequence 2300, Ap  |
| 823 | 6 | 3.3 | 329 | 15 | US-10-389-566-1173   | Sequence 1173, Ap  | 896 | 6 | 3.3 | 352 | 16 | US-10-437-963-181115 | Sequence 181115, A |
| 824 | 6 | 3.3 | 329 | 15 | US-10-389-699-2      | Sequence 2, Appli  | 897 | 6 | 3.3 | 352 | 16 | US-10-437-963-161181 | Sequence 161181, A |
| 825 | 6 | 3.3 | 330 | 16 | US-10-425-114-52140  | Sequence 52140, A  | 898 | 6 | 3.3 | 354 | 15 | US-10-425-114-67272  | Sequence 67272, A  |
| 826 | 6 | 3.3 | 330 | 16 | US-10-437-963-173207 | Sequence 173207, A | 899 | 6 | 3.3 | 354 | 16 | US-10-437-963-170365 | Sequence 170365, A |
| 827 | 6 | 3.3 | 331 | 14 | US-10-233-926-23     | Sequence 23, Appl  | 900 | 6 | 3.3 | 354 | 16 | US-10-437-963-191232 | Sequence 191232, A |
| 828 | 6 | 3.3 | 331 | 15 | US-10-074-978A-177   | Sequence 177, Appl | 901 | 6 | 3.3 | 354 | 16 | US-10-437-963-155475 | Sequence 155475, A |
| 829 | 6 | 3.3 | 331 | 15 | US-10-389-566-1174   | Sequence 1174, Ap  | 902 | 6 | 3.3 | 355 | 16 | US-10-437-963-155475 | Sequence 5, Appli  |
| 830 | 6 | 3.3 | 331 | 15 | US-10-282-122A-60376 | Sequence 60376, A  | 903 | 6 | 3.3 | 356 | 14 | US-09-931-733-5      | Sequence 5, Appli  |
| 831 | 6 | 3.3 | 332 | 15 | US-10-389-566-491    | Sequence 491, Appl | 904 | 6 | 3.3 | 356 | 14 | US-10-177-293-214    | Sequence 214, App  |
| 832 | 6 | 3.3 | 332 | 15 | US-10-389-566-1816   | Sequence 1816, Ap  | 905 | 6 | 3.3 | 356 | 15 | US-10-231-956A-121   | Sequence 121, App  |
| 833 | 6 | 3.3 | 332 | 16 | US-10-437-963-167902 | Sequence 167902, A | 906 | 6 | 3.3 | 356 | 16 | US-10-437-963-201223 | Sequence 201223, A |
| 834 | 6 | 3.3 | 332 | 16 | US-10-437-963-184955 | Sequence 184955, A | 907 | 6 | 3.3 | 357 | 15 | US-10-282-122A-61145 | Sequence 61145, A  |
| 835 | 6 | 3.3 | 333 | 9  | US-09-917-974-2      | Sequence 2, Appli  | 908 | 6 | 3.3 | 357 | 16 | US-10-437-963-168822 | Sequence 168822, A |
| 836 | 6 | 3.3 | 333 | 13 | US-10-013-056-2      | Sequence 2, Appli  | 909 | 6 | 3.3 | 358 | 16 | US-10-437-963-155519 | Sequence 155519, A |
| 837 | 6 | 3.3 | 333 | 16 | US-10-843-747-2      | Sequence 115741, A | 910 | 6 | 3.3 | 359 | 9  | US-09-712-363-226    | Sequence 226, App  |
| 838 | 6 | 3.3 | 333 | 17 | US-10-843-747-2      | Sequence 2, Appli  | 911 | 6 | 3.3 | 359 | 9  | US-09-738-626-3938   | Sequence 3938, Ap  |
| 839 | 6 | 3.3 | 334 | 9  | US-09-925-300-1574   | Sequence 1574, Ap  | 912 | 6 | 3.3 | 359 | 14 | US-10-369-493-10782  | Sequence 10782, A  |
| 840 | 6 | 3.3 | 334 | 15 | US-10-108-260A-4679  | Sequence 4679, Ap  | 913 | 6 | 3.3 | 359 | 15 | US-10-282-122A-63785 | Sequence 63785, A  |
| 841 | 6 | 3.3 | 335 | 16 | US-10-437-963-171160 | Sequence 171160, A | 914 | 6 | 3.3 | 359 | 15 | US-10-282-122A-64654 | Sequence 64654, A  |
| 842 | 6 | 3.3 | 335 | 16 | US-10-437-963-193518 | Sequence 193518, A | 915 | 6 | 3.3 | 360 | 14 | US-10-156-761-14691  | Sequence 14691, A  |
| 843 | 6 | 3.3 | 336 | 9  | US-09-745-763-17     | Sequence 17, Appl  | 916 | 6 | 3.3 | 360 | 16 | US-10-655-780-2      | Sequence 2, Appli  |
| 844 | 6 | 3.3 | 336 | 9  | US-09-799-777-24     | Sequence 24, Appl  | 917 | 6 | 3.3 | 361 | 15 | US-10-425-114-36906  | Sequence 36906, A  |
| 845 | 6 | 3.3 | 337 | 14 | US-10-144-929-97     | Sequence 97, Appl  | 918 | 6 | 3.3 | 362 | 16 | US-10-437-963-178263 | Sequence 178263, A |
| 846 | 6 | 3.3 | 337 | 14 | US-10-369-493-17819  | Sequence 17819, A  | 919 | 6 | 3.3 | 363 | 15 | US-10-425-114-69839  | Sequence 69839, A  |
| 847 | 6 | 3.3 | 337 | 15 | US-10-144-929-97     | Sequence 97, Appl  | 920 | 6 | 3.3 | 363 | 16 | US-10-437-963-159758 | Sequence 159758, A |
| 848 | 6 | 3.3 | 337 | 15 | US-10-282-122A-68028 | Sequence 68028, A  | 921 | 6 | 3.3 | 363 | 16 | US-10-437-963-174398 | Sequence 174398, A |
| 849 | 6 | 3.3 | 338 | 13 | US-10-087-192-315    | Sequence 315, App  | 922 | 6 | 3.3 | 363 | 16 | US-10-437-963-186441 | Sequence 186441, A |
| 850 | 6 | 3.3 | 339 | 16 | US-10-437-963-186333 | Sequence 186333, A | 923 | 6 | 3.3 | 364 | 16 | US-10-437-963-120895 | Sequence 120895, A |
| 851 | 6 | 3.3 | 340 | 9  | US-09-815-242-10726  | Sequence 10726, A  | 924 | 6 | 3.3 | 364 | 16 | US-10-322-281-720    | Sequence 720, App  |
| 852 | 6 | 3.3 | 340 | 16 | US-10-437-963-161248 | Sequence 161248, A | 925 | 6 | 3.3 | 365 | 15 | US-10-282-122A-55983 | Sequence 55983, A  |
| 853 | 6 | 3.3 | 340 | 16 | US-10-437-963-174295 | Sequence 174295, A | 926 | 6 | 3.3 | 365 | 16 | US-10-437-963-114548 | Sequence 114548, A |
| 854 | 6 | 3.3 | 341 | 9  | US-09-738-626-5818   | Sequence 5818, Ap  | 927 | 6 | 3.3 | 365 | 16 | US-10-437-963-140350 | Sequence 140350, A |
| 855 | 6 | 3.3 | 342 | 14 | US-10-233-926-18     | Sequence 18, Appl  | 928 | 6 | 3.3 | 365 | 16 | US-10-437-963-150586 | Sequence 150586, A |
| 856 | 6 | 3.3 | 343 | 14 | US-10-369-493-11303  | Sequence 11303, A  | 929 | 6 | 3.3 | 366 | 16 | US-10-437-963-148650 | Sequence 148650, A |
| 857 | 6 | 3.3 | 343 | 16 | US-10-437-963-114052 | Sequence 114052, A | 930 | 6 | 3.3 | 368 | 16 | US-10-437-963-127150 | Sequence 127150, A |
| 858 | 6 | 3.3 | 343 | 16 | US-10-437-963-184162 | Sequence 184162, A | 931 | 6 | 3.3 | 368 | 16 | US-10-437-963-144630 | Sequence 144630, A |
| 859 | 6 | 3.3 | 344 | 9  | US-09-086-118-27     | Sequence 184162, A | 932 | 6 | 3.3 | 368 | 16 | US-10-437-963-192978 | Sequence 192978, A |
| 860 | 6 | 3.3 | 344 | 11 | US-09-809-665A-101   | Sequence 101, App  | 933 | 6 | 3.3 | 368 | 16 | US-10-767-701-38116  | Sequence 38116, A  |
| 861 | 6 | 3.3 | 344 | 14 | US-10-034-158-7      | Sequence 7, Appl   | 934 | 6 | 3.3 | 369 | 14 | US-10-156-761-11423  | Sequence 11423, A  |
| 862 | 6 | 3.3 | 344 | 14 | US-10-010-390-11     | Sequence 11, Appl  | 935 | 6 | 3.3 | 369 | 14 | US-10-156-761-14259  | Sequence 14259, A  |
| 863 | 6 | 3.3 | 344 | 9  | US-09-880-371-11     | Sequence 11, Appl  | 936 | 6 | 3.3 | 370 | 15 | US-10-282-122A-64566 | Sequence 64566, A  |
| 864 | 6 | 3.3 | 344 | 9  | US-09-879-248-15     | Sequence 15, Appl  | 937 | 6 | 3.3 | 370 | 15 | US-10-437-963-182244 | Sequence 182244, A |
| 865 | 6 | 3.3 | 344 | 9  | US-09-770-693-7      | Sequence 7, Appli  | 938 | 6 | 3.3 | 370 | 16 | US-10-437-963-190966 | Sequence 190966, A |
| 866 | 6 | 3.3 | 344 | 9  | US-09-766-348-7      | Sequence 7, Appli  | 939 | 6 | 3.3 | 371 | 16 | US-10-437-963-128225 | Sequence 128225, A |
| 867 | 6 | 3.3 | 344 | 11 | US-09-809-665A-101   | Sequence 101, App  | 940 | 6 | 3.3 | 371 | 16 | US-10-437-963-139404 | Sequence 139404, A |
| 868 | 6 | 3.3 | 344 | 14 | US-10-034-158-7      | Sequence 7, Appl   | 941 | 6 | 3.3 | 371 | 16 | US-10-437-963-181159 | Sequence 181159, A |
| 869 | 6 | 3.3 | 344 | 14 | US-10-010-390-11     | Sequence 11, Appl  | 942 | 6 | 3.3 | 372 | 14 | US-10-028-072-106    | Sequence 106, App  |
| 870 | 6 | 3.3 | 344 | 14 | US-10-233-926-22     | Sequence 22, Appl  | 943 | 6 | 3.3 | 372 | 14 | US-10-140-808-106    | Sequence 106, App  |
| 871 | 6 | 3.3 | 344 | 14 | US-10-017-161-1954   | Sequence 1954, Ap  | 944 | 6 | 3.3 | 372 | 14 | US-10-131-049-106    | Sequence 106, App  |
| 872 | 6 | 3.3 | 344 | 14 | US-10-387-806-27     | Sequence 27, Appl  | 945 | 6 | 3.3 | 372 | 14 | US-10-123-904-106    | Sequence 106, App  |
| 873 | 6 | 3.3 | 344 | 14 | US-10-292-798-1602   | Sequence 1602, Ap  | 946 | 6 | 3.3 | 372 | 14 | US-10-140-470-106    | Sequence 106, App  |
| 874 | 6 | 3.3 | 344 | 15 | US-10-389-566-709    | Sequence 709, App  | 947 | 6 | 3.3 | 372 | 14 | US-10-175-746-106    | Sequence 106, App  |
| 875 | 6 | 3.3 | 344 | 15 | US-10-282-122A-67153 | Sequence 67153, A  | 948 | 6 | 3.3 | 372 | 14 | US-10-143-114-106    | Sequence 106, App  |
| 876 | 6 | 3.3 | 344 | 15 | US-10-425-114-46057  | Sequence 46057, A  | 949 | 6 | 3.3 | 372 | 14 | US-10-140-002-106    | Sequence 106, App  |
| 877 | 6 | 3.3 | 345 | 16 | US-10-437-963-109343 | Sequence 109343, A | 950 | 6 | 3.3 | 372 | 14 | US-10-142-419-106    | Sequence 106, App  |
| 878 | 6 | 3.3 | 345 | 16 | US-10-437-963-191264 | Sequence 191264, A | 951 | 6 | 3.3 | 372 | 14 | US-10-142-419-106    | Sequence 106, App  |
| 879 | 6 | 3.3 | 348 | 16 | US-10-437-963-150219 | Sequence 150219, A | 952 | 6 | 3.3 | 372 | 14 | US-10-133-262-106    | Sequence 106, App  |
| 880 | 6 | 3.3 | 349 | 14 | US-10-437-963-172625 | Sequence 172625, A | 953 | 6 | 3.3 | 372 | 14 | US-10-142-431-106    | Sequence 106, App  |
| 881 | 6 | 3.3 | 349 | 14 | US-10-156-761-9652   | Sequence 9652, Ap  | 954 | 6 | 3.3 | 372 | 14 | US-10-142-431-106    | Sequence 106, App  |
| 882 | 6 | 3.3 | 349 | 14 | US-10-156-761-13815  | Sequence 13815, A  | 955 | 6 | 3.3 | 372 | 14 | US-10-140-002-106    | Sequence 106, App  |
| 883 | 6 | 3.3 | 349 | 14 | US-10-233-926-2      | Sequence 2, Appli  | 956 | 6 | 3.3 | 372 | 14 | US-10-142-419-106    | Sequence 106, App  |
| 884 | 6 | 3.3 | 349 | 15 | US-10-424-599-214641 | Sequence 214641, A | 957 | 6 | 3.3 | 372 | 14 | US-10-142-419-106    | Sequence 106, App  |
| 885 | 6 | 3.3 | 350 | 15 | US-10-424-599-190192 | Sequence 190192, A | 958 | 6 | 3.3 | 372 | 14 | US-10-142-431-106    | Sequence 106, App  |
| 886 | 6 | 3.3 | 350 | 16 | US-10-437-963-179869 | Sequence 179869, A | 959 | 6 | 3.3 | 372 | 14 | US-10-142-431-106    | Sequence 106, App  |
| 887 | 6 | 3.3 | 350 | 16 | US-10-437-963-185837 | Sequence 185837, A | 960 | 6 | 3.3 | 372 | 14 | US-10-141-755-106    | Sequence 106, App  |
| 888 | 6 | 3.3 | 351 | 9  | US-09-963-340-2      | Sequence 2, Appli  | 961 | 6 | 3.3 | 372 | 14 | US-10-143-032-106    | Sequence 106, App  |
| 889 | 6 | 3.3 | 351 | 14 | US-10-356-076-2      | Sequence 2, Appli  | 962 | 6 | 3.3 | 372 | 14 | US-10-123-108-106    | Sequence 106, App  |





; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 117523  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_2091C.1.pep  
US-10-437-963-117523

Query Match 4.4%; Score 8; DB 16; Length 201;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 PLPLGGGG 170  
|||  
Db 22 PLPLGGGG 29

RESULT 5  
US-10-425-114-39062  
; Sequence 39062, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 39062  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3732-055-F11\_FLI.pep  
US-10-425-114-39062

Query Match 4.4%; Score 8; DB 15; Length 244;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 PLPLGGGG 170  
|||  
Db 145 PLPLGGGG 152

RESULT 6  
US-10-437-963-164580  
; Sequence 164580, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 164580  
; LENGTH: 369  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_63465C.1.pep  
US-10-437-963-164580

Query Match 4.4%; Score 8; DB 16; Length 369;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172  
|||  
Db 44 PLGGGGAR 51

RESULT 7  
US-10-369-493-2997  
; Sequence 2997, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2997  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: Thermotoga maritima  
US-10-369-493-2997

Query Match 4.4%; Score 8; DB 14; Length 476;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 RLTVKEIK 146  
|||  
Db 25 RLTVKEIK 32

RESULT 8  
US-09-826-509-521  
; Sequence 521, Application US/09826509  
; Publication No. US20030204073A1  
; GENERAL INFORMATION:  
; APPLICANT: Lehmann-Brunnsma, Karin  
; APPLICANT: Liaw, Chen W.  
; APPLICANT: Lin, I-Lin  
; TITLE OF INVENTION: No. US20030204073A1-Endogenous, Constitutively Activated Known G  
; TITLE OF INVENTION: Protein-Coupled Receptors

```
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 521
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-521

Query Match      4.4%; Score 8; DB 10; Length 532;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      122 RLVVKADG 129
Db      382 RLVVKADG 389
      |||||

RESULT 9
US-10-225-928-2
; Sequence 2, Application US/10225928
; Publication No. US20030092041A1
; GENERAL INFORMATION:
; APPLICANT: White, David
; TITLE OF INVENTION: NOVEL USE FOR MUSCARINIC RECEPTOR M5 IN
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF METABOLIC DISORDERS
; FILE REFERENCE: MP101-210P1RM
; CURRENT APPLICATION NUMBER: US/10/225,928
; CURRENT FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 532
; TYPE: PRT
; ORGANISM: human
US-10-225-928-2

Query Match      4.4%; Score 8; DB 14; Length 532;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      122 RLVVKADG 129
Db      382 RLVVKADG 389
      |||||

RESULT 10
US-10-225-567A-195
; Sequence 195, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burnex, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 195
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-10-225-567A-195

Query Match      4.4%; Score 8; DB 14; Length 532;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      122 RLVVKADG 129
Db      382 RLVVKADG 389
      |||||

RESULT 11
US-10-366-288-34
; Sequence 34, Application US/10366288
; Publication No. US20030216288A1
; GENERAL INFORMATION:
; APPLICANT: Powell, Douglas S.
; APPLICANT: Welch, Nadine S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING 1414, 1481, 1553,
; TITLE OF INVENTION: 34021, 1720, 1683, 1552, 1682, 1675, 12825, 9952, 5816,
; TITLE OF INVENTION: 10002, 1611, 1371, 14324, 126, 270, 312, 167, 326, 18926,
; TITLE OF INVENTION: 6747, 1793, 1784 OR 2045 MOLECULES
; FILE REFERENCE: MP102-025P1RMNMIM
; CURRENT APPLICATION NUMBER: US/10/366,288
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 60/357,391
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/380,249
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 60/391,306
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/406,297
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 60/412,007
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 60/417,508
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/432,318
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-366-288-34

Query Match      4.4%; Score 8; DB 14; Length 532;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      122 RLVVKADG 129
Db      382 RLVVKADG 389
      |||||

RESULT 12
US-10-280-066-48
; Sequence 48, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Brissette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING TARGET BINDI
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
```

; PRIOR APPLICATION NUMBER: 60/345,471  
; PRIOR FILING DATE: 2001-10-24  
; NUMBER OF SEQ ID NOS: 537  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 48  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; NAME/KEY: MISC FEATURE  
; FEATURE:  
; OTHER INFORMATION: DGI-2-20F-PP-E11  
US-10-280-066-48

Query Match 3.9%; Score 7; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGAR 172  
Db 11 LGGGGAR 17

RESULT 13  
US-10-609-217-1048  
; Sequence 1048, Application US/10609217  
; Publication No. US20040044188A1  
; GENERAL INFORMATION:  
; APPLICANT: FEIGE, ULRICH  
; APPLICANT: LIU, CHUAN-FA  
; APPLICANT: CHEETHAM, JANET C.  
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS  
; FILE REFERENCE: A-527  
; CURRENT APPLICATION NUMBER: US/10/609,217  
; PRIOR FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: US/09/428,082B  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/105,371  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 1133  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1048  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IL-1 ANTAGONIST  
; NAME/KEY: misc.feature  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Fc domain attached at Position 20 of the C-terminus  
US-10-609-217-1048

Query Match 3.9%; Score 7; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPLGGGG 170  
Db 13 LPLGGGG 19

RESULT 14  
US-10-632-388-1048  
; Sequence 1048, Application US/10632388  
; Publication No. US20040053845A1  
; GENERAL INFORMATION:  
; APPLICANT: FEIGE, ULRICH  
; APPLICANT: LIU, CHUAN-FA  
; APPLICANT: CHEETHAM, JANET C.  
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS  
; FILE REFERENCE: A-527

; CURRENT APPLICATION NUMBER: US/10/632,388  
; CURRENT FILING DATE: 2003-07-31  
; PRIOR APPLICATION NUMBER: US/09/428,082B  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/105,371  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 1133  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1048  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IL-1 ANTAGONIST  
; NAME/KEY: misc.feature  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Fc domain attached at Position 20 of the C-terminus  
US-10-632-388-1048

Query Match 3.9%; Score 7; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPLGGGG 170  
Db 13 LPLGGGG 19

RESULT 15  
US-10-651-723-1048  
; Sequence 1048, Application US/10651723  
; Publication No. US20040057953A1  
; GENERAL INFORMATION:  
; APPLICANT: FEIGE, ULRICH  
; APPLICANT: LIU, CHUAN-FA  
; APPLICANT: CHEETHAM, JANET C.  
; APPLICANT: BOONE, THOMAS CHARLES  
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS  
; FILE REFERENCE: A-527  
; CURRENT APPLICATION NUMBER: US/10/651,723  
; CURRENT FILING DATE: 2003-08-29  
; PRIOR APPLICATION NUMBER: US/09/428,082B  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/105,371  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 1133  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1048  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IL-1 ANTAGONIST  
; NAME/KEY: misc.feature  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Fc domain attached at Position 20 of the C-terminus  
US-10-651-723-1048

Query Match 3.9%; Score 7; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPLGGGG 170  
Db 13 LPLGGGG 19

RESULT 16  
US-10-645-761-1048  
; Sequence 1048, Application US/10645761  
; Publication No. US20040071712A1

; GENERAL INFORMATION:  
; APPLICANT: FEIGE, ULRICH  
; APPLICANT: LIU, CHUAN-FA  
; APPLICANT: CHEETHAM, JANET C.  
; APPLICANT: BOONE, THOMAS CHARLES  
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS  
; FILE REFERENCE: A-527  
; CURRENT APPLICATION NUMBER: US/10/645,761  
; CURRENT FILING DATE: 2003-08-18  
; PRIOR APPLICATION NUMBER: US/09/428,082B  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/105,371  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 1133  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1048  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IL-1 ANTAGONIST  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Fc domain attached at Position 20 of the C-terminus  
US-10-645-761-1048

Query Match 3.9%; Score 7; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 LPLGGGG 170  
Db 13 LPLGGGG 19  
|||||

RESULT 17  
US-10-666-696-1048  
; Sequence 1048, Application US/10666696  
; Publication No. US20040077022A1  
; GENERAL INFORMATION:  
; APPLICANT: FEIGE, ULRICH  
; APPLICANT: LIU, CHUAN-FA  
; APPLICANT: CHEETHAM, JANET C.  
; APPLICANT: BOONE, THOMAS CHARLES  
; APPLICANT: GUDAS, JEAN MARIE  
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS  
; FILE REFERENCE: A-527A  
; CURRENT APPLICATION NUMBER: US/10/666,696  
; CURRENT FILING DATE: 2003-09-19  
; PRIOR APPLICATION NUMBER: US/09/563,286C  
; PRIOR FILING DATE: 2000-05-03  
; PRIOR APPLICATION NUMBER: 09/428,082  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/105,371  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 1157  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1048  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IL-1 R ANTAGONIST  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Fc domain attached at Position 20 of the C-terminus  
US-10-666-696-1048

Query Match 3.9%; Score 7; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 LPLGGGG 170  
Db 13 LPLGGGG 19  
|||||

RESULT 18  
US-10-653-048-1048  
; Sequence 1048, Application US/10653048  
; Publication No. US20040087778A1  
; GENERAL INFORMATION:  
; APPLICANT: FEIGE, ULRICH  
; APPLICANT: LIU, CHUAN-FA  
; APPLICANT: CHEETHAM, JANET C.  
; APPLICANT: BOONE, THOMAS CHARLES  
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS  
; FILE REFERENCE: A-527  
; CURRENT APPLICATION NUMBER: US/10/653,048  
; CURRENT FILING DATE: 2003-08-23  
; PRIOR APPLICATION NUMBER: US/09/428,082B  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 60/105,371  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 1133  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1048  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: IL-1 ANTAGONIST  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Fc domain attached at Position 20 of the C-terminus  
US-10-653-048-1048

Query Match 3.9%; Score 7; DB 15; Length 20;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 LPLGGGG 170  
Db 13 LPLGGGG 19  
|||||

RESULT 19  
US-10-437-963-163985  
; Sequence 163985, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 163985  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(53)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_6292C.1.pep  
US-10-437-963-163985

Query Match 3.9%; Score 7; DB 16; Length 53;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 LTVKEIK 146  
|||  
Db 6 LTVKEIK 12

## RESULT 20

US-10-767-701-50522  
; Sequence 50522, Application US/10767701  
; Publication No. US20040172664A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT FILING DATE: 2004-01-29  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 50522  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3480-013-P1-K1-E5.pep  
US-10-767-701-50522

Query Match 3.9%; Score 7; DB 16; Length 69;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGAR 172  
|||  
Db 60 LGGGGAR 66

## RESULT 21

US-10-437-963-140344  
; Sequence 140344, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT FILING DATE: 2003-05-14  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 140344  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_41550C.1.pep  
US-10-437-963-140344

Query Match 3.9%; Score 7; DB 16; Length 85;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGSCGPA 84  
|||  
Db 71 GGSCGPA 77

## RESULT 22

US-10-425-114-53565  
; Sequence 53565, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT FILING DATE: 2003-04-28  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 53565  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3079-055-A7\_FLI.pep  
US-10-425-114-53565

Query Match 3.9%; Score 7; DB 15; Length 96;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 AEHDGHL 90  
|||  
Db 82 AEHDGHL 88

## RESULT 23

US-10-424-599-158607  
; Sequence 158607, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT FILING DATE: 2003-04-28  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 158607  
; LENGTH: 105  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_114240C.1.pep  
US-10-424-599-158607

Query Match 3.9%; Score 7; DB 15; Length 105;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LPLGGGG 170  
|||  
Db 28 LPLGGGG 34

## RESULT 24

US-10-425-114-40171  
; Sequence 40171, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 40171  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701052313\_FJI.pep  
US-10-425-114-40171

Query Match 3.9%; Score 7; DB 15; Length 107;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 163 PLPLGGG 169  
Db 58 PLPLGGG 64  
|||||

RESULT 25  
US-10-424-599-261961  
; Sequence 261961, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 261961  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(108)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_78573C.1.pep  
US-10-424-599-261961

Query Match 3.9%; Score 7; DB 15; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LFFVTSI 11  
Db 11 LFFVTSI 17  
|||||

RESULT 26  
US-10-437-963-168698  
; Sequence 168698, Application US/10437963  
; Publication No. US20040123343A1

; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 168698  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(109)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_67188C.1.pep  
US-10-437-963-168698

Query Match 3.9%; Score 7; DB 16; Length 109;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 GDLPLV 124  
Db 21 GDLPLV 27  
|||||

RESULT 27  
US-10-424-599-160841  
; Sequence 160841, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 160841  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(118)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_116257C.1.pep  
US-10-424-599-160841

Query Match 3.9%; Score 7; DB 15; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 163 PLPLGGG 169  
Db 55 PLPLGGG 61  
|||||

RESULT 28

```
US-10-437-963-190020
; Sequence 190020, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihwei
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190020
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(121)
; FEATURE:
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86473C.1.pap
US-10-437-963-190020

Query Match          3.9%; Score 7; DB 16; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 DNGIKQS 44
Db 59 DNGIKQS 65

RESULT 29
US-10-282-122A-57872
; Sequence 57872, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57138
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
```

```
US-10-282-122A-57872
; Sequence 57872, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57138
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Enterococcus faecalis

Query Match          3.9%; Score 7; DB 15; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 TDLKGLP 67
Db 95 TDLKGLP 101

RESULT 30
US-10-282-122A-57138
; Sequence 57138, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57138
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
```

US-10-282-122A-57138

Query Match 3.9%; Score 7; DB 15; Length 124;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 TDLKGLP 67  
|||||||  
Db 95 TDLKGLP 101

RESULT 31

US-10-767-701-40863  
; Sequence 40863, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 40863  
; LENGTH: 124  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C10838\_1.pep

Query Match 3.9%; Score 7; DB 15; Length 124;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 LGGGGAR 172  
|||||||  
Db 13 LGGGGAR 19

RESULT 32

US-10-767-701-58300  
; Sequence 58300, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 58300  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 30978261.pep

Query Match 3.9%; Score 7; DB 15; Length 140;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 GGGGARI 173  
|||||||  
Db 11 GGGGARI 17

RESULT 33

US-10-437-963-144207  
; Sequence 144207, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 144207  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(141)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_45043C.1.pep

US-10-437-963-144207

Query Match 3.9%; Score 7; DB 16; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 GGGGARI 173  
|||||||  
Db 66 GGGGARI 72

RESULT 34

US-10-767-701-37046  
; Sequence 37046, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 37046  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(141)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C14213\_1.pep

Query Match 3.9%; Score 7; DB 16; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 LGGGGAR 172  
|||||||  
Db 98 LGGGGAR 104



```
RESULT 35
US-10-437-963-170400
; Sequence 170400, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170400
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(144)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68730C.1.pep
US-10-437-963-170400

Query Match          3.9%; Score 7; DB 16; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGAR 172
DB 130 LGGGGAR 136

RESULT 36
US-10-767-701-40660
; Sequence 40660, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40660
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C7960_1.pep
US-10-767-701-40660

Query Match          3.9%; Score 7; DB 16; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 PLPLGGG 169
DB 78 PLPLGGG 84

RESULT 37
US-10-437-963-167300
; Sequence 167300, Application US/10437963
; Publication No. US20040123343A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 167300
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(144)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_65924C.1.pep
US-10-437-963-167300

Query Match          3.9%; Score 7; DB 16; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GHLTAGL 94
DB 11 GHLTAGL 17

RESULT 38
US-10-424-599-269534
; Sequence 269534, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 269534
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85409C.1.pep
US-10-424-599-269534

Query Match          3.9%; Score 7; DB 15; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LKGLPAG 69
DB 36 LKGLPAG 42

RESULT 39
US-10-104-047-2908
; Sequence 2908, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
```

```
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2908
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2908

Query Match          3.9%; Score 7; DB 14; Length 157;
Best Local Similarity 100.0%; Pred.No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      165 PLGGGGA 171
Db      34 PLGGGGA 40
      |||||
      |||||

RESULT 40
US-10-424-599-268560
; Sequence 268560, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 268560
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84532C.1.pep
US-10-424-599-268560

Query Match          3.9%; Score 7; DB 15; Length 166;
Best Local Similarity 100.0%; Pred.No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      71 HGFHIHE 77
Db      49 HGFHIHE 55
      |||||
      |||||

Search completed: October 26, 2004, 10:14:50
Job time : 144 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 09:54:25 ; Search time 39 Seconds  
(without alignments)  
444.077 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 180

Sequence: 1 MKIKLFVTSIVTISLTSI.....DKPLPLGGGARIACGVIPN 180

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 11    | 6.1         | 201    | A10409   | superoxide dismuta |
| 2          | 9     | 5.0         | 154    | A33893   | superoxide dismuta |
| 3          | 9     | 5.0         | 171    | A82183   | superoxide dismuta |
| 4          | 9     | 5.0         | 174    | A23582   | superoxide dismuta |
| 5          | 8     | 4.4         | 87     | I39485   | superoxide dismuta |
| 6          | 8     | 4.4         | 173    | 1 DSFOCL | superoxide dismuta |
| 7          | 8     | 4.4         | 173    | 2 GS5771 | superoxide dismuta |
| 8          | 8     | 4.4         | 173    | 2 JC6004 | superoxide dismuta |
| 9          | 8     | 4.4         | 173    | 2 AF0694 | copper-zinc supero |
| 10         | 8     | 4.4         | 173    | 2 C90923 | superoxide dismuta |
| 11         | 8     | 4.4         | 175    | 2 F90768 | superoxide dismuta |
| 12         | 8     | 4.4         | 175    | 2 F90877 | probable copper/zi |
| 13         | 8     | 4.4         | 187    | 1 A41654 | probable copper/zi |
| 14         | 8     | 4.4         | 187    | 1 B41654 | superoxide dismuta |
| 15         | 8     | 4.4         | 199    | 2 JC5718 | superoxide dismuta |
| 16         | 8     | 4.4         | 274    | 2 F85741 | superoxide dismuta |
| 17         | 8     | 4.4         | 328    | 2 B85842 | hypothetical prote |
| 18         | 8     | 4.4         | 476    | 2 C72299 | probable superoxid |
| 19         | 8     | 4.4         | 531    | 2 JT0531 | sugar kinase - The |
| 20         | 8     | 4.4         | 532    | 2 JT0530 | muscarinic acetyl  |
| 21         | 7     | 3.9         | 129    | 2 AC1716 | glyoxalase I homol |
| 22         | 7     | 3.9         | 129    | 2 AH1345 | glyoxalase I homol |
| 23         | 7     | 3.9         | 140    | 2 F95926 | hypothetical expor |
| 24         | 7     | 3.9         | 171    | 2 F70321 | superoxide dismuta |
| 25         | 7     | 3.9         | 176    | 2 JE0097 | superoxide dismuta |
| 26         | 7     | 3.9         | 184    | 2 S40984 | superoxide dismuta |
| 27         | 7     | 3.9         | 186    | 2 E81855 | superoxide dismuta |
| 28         | 7     | 3.9         | 186    | 2 F81088 | superoxide dismuta |
| 29         | 7     | 3.9         | 221    | 2 JE0098 | superoxide dismuta |

|     |   |     |      |   |        |                     |
|-----|---|-----|------|---|--------|---------------------|
| 30  | 7 | 3.9 | 227  | 2 | F90626 | ATP synthase F0 ch  |
| 31  | 7 | 3.9 | 232  | 2 | AE2049 | ATP-binding protei  |
| 32  | 7 | 3.9 | 238  | 2 | T30864 | stbB protein - Sal  |
| 33  | 7 | 3.9 | 270  | 2 | S45407 | probable membrane   |
| 34  | 7 | 3.9 | 271  | 2 | F86171 | hypothetical prote  |
| 35  | 7 | 3.9 | 308  | 2 | AE2661 | conserved hypothet  |
| 36  | 7 | 3.9 | 311  | 2 | C97443 | hypothetical prote  |
| 37  | 7 | 3.9 | 341  | 2 | E75374 | conserved hypothet  |
| 38  | 7 | 3.9 | 347  | 2 | E90057 | hypothetical prote  |
| 39  | 7 | 3.9 | 382  | 2 | S78285 | conserved hypothet  |
| 40  | 7 | 3.9 | 383  | 2 | C87478 | hypothetical prote  |
| 41  | 7 | 3.9 | 383  | 2 | AG3312 | X-Pro dipeptidase   |
| 42  | 7 | 3.9 | 406  | 2 | T06805 | Rti17-1 protein hom |
| 43  | 7 | 3.9 | 444  | 2 | T03566 | probable trigger f  |
| 44  | 7 | 3.9 | 455  | 2 | A40163 | glycylpeptide N-te  |
| 45  | 7 | 3.9 | 462  | 2 | T46205 | hypothetical prote  |
| 46  | 7 | 3.9 | 490  | 2 | T05444 | hypothetical prote  |
| 47  | 7 | 3.9 | 492  | 2 | H85783 | probable aldehyde   |
| 48  | 7 | 3.9 | 492  | 2 | D90935 | probable aldehyde   |
| 49  | 7 | 3.9 | 555  | 2 | E87003 | probable DNA methy  |
| 50  | 7 | 3.9 | 558  | 2 | T47505 | hypothetical prote  |
| 51  | 7 | 3.9 | 570  | 2 | D96766 | protein glucosyltr  |
| 52  | 7 | 3.9 | 578  | 2 | D95888 | hypothetical prote  |
| 53  | 7 | 3.9 | 592  | 2 | T46657 | heat shock protein  |
| 54  | 7 | 3.9 | 667  | 2 | S60709 | alpha 1,2 mannosid  |
| 55  | 7 | 3.9 | 688  | 2 | T18263 | S-layer protein -   |
| 56  | 7 | 3.9 | 756  | 2 | AB1088 | chitinase B homolo  |
| 57  | 7 | 3.9 | 756  | 2 | AB1452 | chitinase B homolo  |
| 58  | 7 | 3.9 | 1048 | 2 | A70592 | hypothetical prote  |
| 59  | 7 | 3.9 | 1073 | 1 | OYHUHX | heat-stable entero  |
| 60  | 7 | 3.9 | 1075 | 1 | OYRTHX | heat-stable entero  |
| 61  | 7 | 3.9 | 1188 | 2 | S50434 | hypothetical prote  |
| 62  | 7 | 3.9 | 1316 | 2 | T50444 | hypothetical prote  |
| 63  | 7 | 3.9 | 1351 | 2 | S44665 | hypothetical UFP00  |
| 64  | 7 | 3.9 | 1421 | 2 | T49500 | ZK370.4 protein -   |
| 65  | 7 | 3.9 | 1487 | 2 | F82083 | hypothetical prote  |
| 66  | 7 | 3.9 | 1679 | 2 | S49802 | glutamate synthase  |
| 67  | 7 | 3.9 | 8243 | 2 | T31307 | probable membrane   |
| 68  | 6 | 3.3 | 45   | 2 | E70236 | type I fatty acid   |
| 69  | 6 | 3.3 | 50   | 2 | S29149 | hypothetical prote  |
| 70  | 6 | 3.3 | 67   | 2 | F72856 | superoxide dismuta  |
| 71  | 6 | 3.3 | 68   | 2 | AC1093 | hypothetical prote  |
| 72  | 6 | 3.3 | 69   | 2 | A10473 | hypothetical prote  |
| 73  | 6 | 3.3 | 72   | 2 | AH3281 | hypothetical cytos  |
| 74  | 6 | 3.3 | 80   | 2 | S29828 | hypothetical prote  |
| 75  | 6 | 3.3 | 82   | 2 | H59101 | cytochrome c553 -   |
| 76  | 6 | 3.3 | 84   | 2 | A43502 | hypothetical prote  |
| 77  | 6 | 3.3 | 95   | 2 | E81137 | MEP70 protein - My  |
| 78  | 6 | 3.3 | 97   | 2 | F97780 | probable membrane   |
| 79  | 6 | 3.3 | 103  | 2 | T29342 | hypothetical prote  |
| 80  | 6 | 3.3 | 104  | 2 | S00120 | hypothetical prote  |
| 81  | 6 | 3.3 | 105  | 2 | D64031 | hypothetical prote  |
| 82  | 6 | 3.3 | 107  | 2 | F81747 | hypothetical prote  |
| 83  | 6 | 3.3 | 107  | 2 | H72764 | hypothetical prote  |
| 84  | 6 | 3.3 | 108  | 2 | C85256 | hypothetical prote  |
| 85  | 6 | 3.3 | 110  | 2 | G72597 | Ribosomal protein   |
| 86  | 6 | 3.3 | 112  | 2 | D95992 | hypothetical prote  |
| 87  | 6 | 3.3 | 114  | 2 | T41969 | hypothetical prote  |
| 88  | 6 | 3.3 | 118  | 2 | T17901 | hypothetical prote  |
| 89  | 6 | 3.3 | 124  | 1 | A70343 | hypothetical prote  |
| 90  | 6 | 3.3 | 126  | 2 | D72570 | aspartate 1-decarb  |
| 91  | 6 | 3.3 | 127  | 2 | AG1425 | hypothetical prote  |
| 92  | 6 | 3.3 | 128  | 2 | C85135 | Ribosomal protein   |
| 93  | 6 | 3.3 | 129  | 2 | B70832 | hypothetical prote  |
| 94  | 6 | 3.3 | 132  | 2 | E72603 | hypothetical prote  |
| 95  | 6 | 3.3 | 133  | 2 | S73769 | adenine phosphorib  |
| 96  | 6 | 3.3 | 133  | 2 | S10038 | hypothetical prote  |
| 97  | 6 | 3.3 | 135  | 2 | T51240 | scarecrow-like pro  |
| 98  | 6 | 3.3 | 135  | 2 | AF3551 | hypothetical prote  |
| 99  | 6 | 3.3 | 136  | 2 | D82982 | cytochrome c5 PAS3  |
| 100 | 6 | 3.3 | 137  | 2 | S29042 | Na+-transporting A  |
| 101 | 6 | 3.3 | 137  | 2 | A03865 | hypothetical prote  |
| 102 | 6 | 3.3 | 137  | 2 | A03865 | hypothetical prote  |

|     |   |     |     |   |        |                    |     |   |     |     |   |        |                     |
|-----|---|-----|-----|---|--------|--------------------|-----|---|-----|-----|---|--------|---------------------|
| 103 | 6 | 3.3 | 139 | 2 | T09309 | immediate-early pr | 176 | 6 | 3.3 | 197 | 2 | F75498 | conserved hypotet   |
| 104 | 6 | 3.3 | 143 | 2 | B85581 | unknown protein en | 177 | 6 | 3.3 | 201 | 2 | AG0442 | probable phosphos   |
| 105 | 6 | 3.3 | 143 | 2 | T40617 | hypothetical prote | 178 | 6 | 3.3 | 203 | 2 | B71411 | hypothetical prote  |
| 106 | 6 | 3.3 | 145 | 2 | C70328 | conserved hypotet  | 179 | 6 | 3.3 | 203 | 2 | T25609 | hypothetical prote  |
| 107 | 6 | 3.3 | 146 | 2 | E87147 | SecB preprotein tr | 180 | 6 | 3.3 | 205 | 2 | C86862 | guanylate kinase (  |
| 108 | 6 | 3.3 | 150 | 2 | S05021 | superoxide dismuta | 181 | 6 | 3.3 | 207 | 2 | F87470 | transcription regu  |
| 109 | 6 | 3.3 | 150 | 2 | T00195 | hypothetical prote | 182 | 6 | 3.3 | 208 | 1 | A39432 | heparin-binding EG  |
| 110 | 6 | 3.3 | 151 | 1 | DSWFCZ | superoxide dismuta | 183 | 6 | 3.3 | 217 | 2 | H72390 | dihydroorotate deh  |
| 111 | 6 | 3.3 | 151 | 1 | S09568 | superoxide dismuta | 184 | 6 | 3.3 | 219 | 1 | F69500 | riboflavin-specifi  |
| 112 | 6 | 3.3 | 151 | 2 | C84514 | hypothetical prote | 185 | 6 | 3.3 | 220 | 2 | D85991 | probable transcrip  |
| 113 | 6 | 3.3 | 152 | 2 | S22508 | superoxide dismuta | 186 | 6 | 3.3 | 220 | 2 | B85119 | potential acrf/en   |
| 114 | 6 | 3.3 | 152 | 2 | T06570 | superoxide dismuta | 187 | 6 | 3.3 | 220 | 2 | H91145 | probable transcrip  |
| 115 | 6 | 3.3 | 153 | 2 | F83481 | probable HIR famil | 188 | 6 | 3.3 | 221 | 1 | Q1BP1L | multiple specifici  |
| 116 | 6 | 3.3 | 153 | 2 | G72583 | hypothetical prote | 189 | 6 | 3.3 | 221 | 2 | F85910 | multiple specifici  |
| 117 | 6 | 3.3 | 154 | 1 | DSBYC  | superoxide dismuta | 190 | 6 | 3.3 | 221 | 2 | F91066 | multiple specifici  |
| 118 | 6 | 3.3 | 154 | 2 | JC1192 | superoxide dismuta | 191 | 6 | 3.3 | 221 | 2 | E90730 | multiple specifici  |
| 119 | 6 | 3.3 | 155 | 2 | C86206 | hypothetical prote | 192 | 6 | 3.3 | 222 | 2 | A81084 | O-methyltransfera   |
| 120 | 6 | 3.3 | 156 | 2 | T12204 | superoxide dismuta | 193 | 6 | 3.3 | 222 | 2 | A81860 | probable methyltra  |
| 121 | 6 | 3.3 | 157 | 2 | T08157 | ribosomal protein  | 194 | 6 | 3.3 | 223 | 2 | H71485 | hypothetical prote  |
| 122 | 6 | 3.3 | 158 | 2 | S18743 | superoxide dismuta | 195 | 6 | 3.3 | 224 | 2 | A81665 | ribosomal protein   |
| 123 | 6 | 3.3 | 158 | 2 | S57779 | oleosin 2 - barley | 196 | 6 | 3.3 | 224 | 2 | H71506 | probable S3 riboso  |
| 124 | 6 | 3.3 | 159 | 2 | T48956 | hypothetical prote | 197 | 6 | 3.3 | 226 | 2 | T46058 | porin-like protein  |
| 125 | 6 | 3.3 | 160 | 2 | AG2191 | glucokinase [imp   | 198 | 6 | 3.3 | 230 | 2 | AG3651 | hypothetical prote  |
| 126 | 6 | 3.3 | 160 | 2 | B95973 | hypothetical expor | 199 | 6 | 3.3 | 232 | 2 | E70789 | probable phosphos   |
| 127 | 6 | 3.3 | 161 | 2 | F75542 | conserved hypotet  | 200 | 6 | 3.3 | 233 | 1 | A75174 | probable phosphos   |
| 128 | 6 | 3.3 | 161 | 2 | AE1322 | hypothetical prote | 201 | 6 | 3.3 | 233 | 2 | D86603 | muramidase [import  |
| 129 | 6 | 3.3 | 161 | 2 | C70613 | probable secE - My | 202 | 6 | 3.3 | 233 | 2 | A72022 | muramidase (invasi  |
| 130 | 6 | 3.3 | 161 | 2 | AH2142 | hypothetical prote | 203 | 6 | 3.3 | 233 | 2 | A97861 | zinc ABC transport  |
| 131 | 6 | 3.3 | 162 | 2 | T51731 | superoxide dismuta | 204 | 6 | 3.3 | 233 | 2 | C71645 | ferric transport A  |
| 132 | 6 | 3.3 | 162 | 2 | A90210 | hypothetical prote | 205 | 6 | 3.3 | 235 | 2 | S33448 | resiniferatoxin-bi  |
| 133 | 6 | 3.3 | 164 | 2 | C89854 | hypothetical prote | 206 | 6 | 3.3 | 237 | 1 | S73990 | ribosomal protein   |
| 134 | 6 | 3.3 | 164 | 2 | A81325 | probable signal-tr | 207 | 6 | 3.3 | 237 | 2 | C83380 | probable pilli asse |
| 135 | 6 | 3.3 | 165 | 2 | AD1359 | bacteriophage prot | 208 | 6 | 3.3 | 237 | 2 | H69820 | conserved hypotet   |
| 136 | 6 | 3.3 | 165 | 2 | G81324 | probable signal-tr | 209 | 6 | 3.3 | 244 | 2 | T06660 | hypothetical prote  |
| 137 | 6 | 3.3 | 166 | 2 | S72918 | hypothetical prote | 210 | 6 | 3.3 | 245 | 2 | T04797 | embryo-specific pr  |
| 138 | 6 | 3.3 | 166 | 2 | G83176 | hypothetical prote | 211 | 6 | 3.3 | 246 | 2 | AH1925 | hypothetical prote  |
| 139 | 6 | 3.3 | 167 | 2 | B81398 | probable membrane  | 212 | 6 | 3.3 | 246 | 2 | C97226 | dihydroorotate den  |
| 140 | 6 | 3.3 | 168 | 2 | T26984 | hypothetical prote | 213 | 6 | 3.3 | 247 | 2 | E90399 | exodeoxyribonuclea  |
| 141 | 6 | 3.3 | 168 | 2 | B97776 | hypothetical prote | 214 | 6 | 3.3 | 250 | 2 | T03115 | probable fusion pr  |
| 142 | 6 | 3.3 | 169 | 2 | B70390 | superoxide dismuta | 215 | 6 | 3.3 | 252 | 1 | S01821 | glycine-rich prote  |
| 143 | 6 | 3.3 | 169 | 2 | B75357 | probable septum si | 216 | 6 | 3.3 | 252 | 2 | B83584 | probable molybdenu  |
| 144 | 6 | 3.3 | 173 | 2 | I46197 | cell surface prote | 217 | 6 | 3.3 | 253 | 2 | T52613 | chaperonin 21 prec  |
| 145 | 6 | 3.3 | 173 | 2 | A86400 | protein F17L21.28  | 218 | 6 | 3.3 | 253 | 2 | F72710 | hypothetical prote  |
| 146 | 6 | 3.3 | 174 | 2 | F97240 | peptidyl-prolyl ci | 219 | 6 | 3.3 | 253 | 2 | C83906 | hypothetical prote  |
| 147 | 6 | 3.3 | 174 | 2 | T07736 | probable chaperoni | 220 | 6 | 3.3 | 254 | 2 | F64639 | hypothetical prote  |
| 148 | 6 | 3.3 | 175 | 2 | C84544 | probable gluconoki | 221 | 6 | 3.3 | 254 | 2 | T52122 | chaperonin 10 [imp  |
| 149 | 6 | 3.3 | 175 | 2 | AE1149 | hypothetical prote | 222 | 6 | 3.3 | 255 | 2 | A11374 | conserved hypotet   |
| 150 | 6 | 3.3 | 175 | 2 | AE1508 | hypothetical prote | 223 | 6 | 3.3 | 255 | 2 | AG1744 | conserved hypotet   |
| 151 | 6 | 3.3 | 176 | 2 | D87345 | hypothetical prote | 224 | 6 | 3.3 | 255 | 2 | S35078 | hypothetical prote  |
| 152 | 6 | 3.3 | 177 | 2 | C88115 | protein F53C3.3 [i | 225 | 6 | 3.3 | 255 | 2 | AD1840 | hypothetical prote  |
| 153 | 6 | 3.3 | 177 | 2 | T02232 | protein BX06 - com | 226 | 6 | 3.3 | 256 | 2 | JC4627 | fibroblast growth   |
| 154 | 6 | 3.3 | 178 | 2 | T27860 | probable superoxid | 227 | 6 | 3.3 | 256 | 2 | A70514 | hypothetical glyci  |
| 155 | 6 | 3.3 | 178 | 2 | T42661 | hypothetical prote | 228 | 6 | 3.3 | 256 | 2 | G39845 | dihydroorotate deh  |
| 156 | 6 | 3.3 | 180 | 2 | A83993 | hypothetical prote | 229 | 6 | 3.3 | 257 | 1 | S42271 | H+-transporting tw  |
| 157 | 6 | 3.3 | 182 | 2 | H97067 | superoxide dismuta | 230 | 6 | 3.3 | 257 | 2 | B69012 | probable membrane   |
| 158 | 6 | 3.3 | 182 | 2 | G49804 | hypothetical prote | 231 | 6 | 3.3 | 257 | 2 | G72648 | hypothetical prote  |
| 159 | 6 | 3.3 | 185 | 2 | D96572 | protein F12M16.6 [ | 232 | 6 | 3.3 | 257 | 2 | I40170 | hypothetical prote  |
| 160 | 6 | 3.3 | 186 | 2 | PW0008 | chaperonin 60K be  | 233 | 6 | 3.3 | 257 | 2 | B88700 | protein K02B2.3 [i  |
| 161 | 6 | 3.3 | 186 | 2 | H98320 | probable proteinas | 234 | 6 | 3.3 | 262 | 2 | C86793 | hypothetical prote  |
| 162 | 6 | 3.3 | 186 | 2 | AH3065 | proteinase [import | 235 | 6 | 3.3 | 262 | 2 | S29920 | SalY7R protein - v  |
| 163 | 6 | 3.3 | 191 | 2 | E84831 | hypothetical prote | 236 | 6 | 3.3 | 263 | 2 | T37423 | probable 29.8K pro  |
| 164 | 6 | 3.3 | 192 | 2 | A64098 | molybdopterin-guan | 237 | 6 | 3.3 | 263 | 2 | C42521 | A37R protein - vac  |
| 165 | 6 | 3.3 | 193 | 2 | A48320 | major secreted imm | 238 | 6 | 3.3 | 264 | 2 | S29144 | hypothetical prote  |
| 166 | 6 | 3.3 | 193 | 2 | A37195 | probable secreted  | 239 | 6 | 3.3 | 265 | 2 | D36861 | orf3 protein - Lel  |
| 167 | 6 | 3.3 | 193 | 2 | F70923 | probable mpt70 pro | 240 | 6 | 3.3 | 265 | 2 | F72618 | hypothetical prote  |
| 168 | 6 | 3.3 | 194 | 1 | S49184 | phosphinothricin N | 241 | 6 | 3.3 | 265 | 2 | T29322 | hypothetical prote  |
| 169 | 6 | 3.3 | 195 | 2 | C75490 | hypothetical prote | 242 | 6 | 3.3 | 266 | 2 | A69358 | carboxylesterase (  |
| 170 | 6 | 3.3 | 196 | 2 | G72063 | 15 kDa cysteine-ri | 243 | 6 | 3.3 | 266 | 2 | C71027 | probable erythrocy  |
| 171 | 6 | 3.3 | 196 | 2 | H86559 | 15 kDa Cysteine-ri | 244 | 6 | 3.3 | 267 | 2 | E71876 | hypothetical prote  |
| 172 | 6 | 3.3 | 196 | 2 | A95890 | hypothetical prote | 245 | 6 | 3.3 | 267 | 2 | H72590 | hypothetical prote  |
| 173 | 6 | 3.3 | 197 | 2 | G65139 | hypothetical prote | 246 | 6 | 3.3 | 267 | 2 | G83640 | tryptophan synthas  |
| 174 | 6 | 3.3 | 197 | 2 | G91163 | hypothetical prote | 247 | 6 | 3.3 | 268 | 2 | T46755 | membrane protein a  |
| 175 | 6 | 3.3 | 197 | 2 | H86009 | hypothetical prote | 248 | 6 | 3.3 | 268 | 2 | G95253 | zinc ABC transport  |

|     |   |     |                      |     |        |         |     |        |     |   |     |     |   |         |                     |
|-----|---|-----|----------------------|-----|--------|---------|-----|--------|-----|---|-----|-----|---|---------|---------------------|
| 249 | 6 | 3.3 | high-affinity zinc   | 2   | AD1094 | 268     | 2   | AD1094 | 322 | 6 | 3.3 | 327 | 2 | A44168  | homeotic protein    |
| 250 | 6 | 3.3 | high-affinity zinc   | 2   | AB1457 | 268     | 2   | AB1457 | 323 | 6 | 3.3 | 328 | 1 | A41927  | insulin-like growth |
| 251 | 6 | 3.3 | tryptophan synthas   | 324 | 2      | AB30768 | 324 | 324    | 324 | 6 | 3.3 | 328 | 2 | D81429  | probable membrane   |
| 252 | 6 | 3.3 | tryptophan synthas   | 325 | 2      | AB30768 | 325 | 325    | 325 | 6 | 3.3 | 329 | 2 | D97981  | probable choline-p  |
| 253 | 6 | 3.3 | inositol-1(or 4)-m   | 326 | 2      | AG33638 | 326 | 326    | 326 | 6 | 3.3 | 330 | 2 | D95958  | conserved hypothet  |
| 254 | 6 | 3.3 | acid proteinase ea   | 327 | 2      | S63631  | 327 | 327    | 327 | 6 | 3.3 | 331 | 2 | T07980  | probable choline-p  |
| 255 | 6 | 3.3 | probable snescenc    | 328 | 2      | E84578  | 328 | 328    | 328 | 6 | 3.3 | 331 | 2 | AH3081  | conserved hypothet  |
| 256 | 6 | 3.3 | short-chain alcoho   | 329 | 2      | S34678  | 329 | 329    | 329 | 6 | 3.3 | 331 | 2 | H98204  | hypothetical prote  |
| 257 | 6 | 3.3 | tryptophan synthas   | 330 | 2      | D43664  | 330 | 330    | 330 | 6 | 3.3 | 331 | 2 | F83246  | glucokinase PA3193  |
| 258 | 6 | 3.3 | hypothetical prote   | 331 | 2      | T070737 | 331 | 331    | 331 | 6 | 3.3 | 331 | 2 | AD11103 | conserved hypothet  |
| 259 | 6 | 3.3 | voltage-dependent    | 332 | 2      | T091116 | 332 | 332    | 332 | 6 | 3.3 | 331 | 2 | AD1465  | conserved hypothet  |
| 260 | 6 | 3.3 | 4-amino-4-deoxycho   | 333 | 2      | C82128  | 333 | 333    | 333 | 6 | 3.3 | 332 | 2 | H84730  | probable phospholi  |
| 261 | 6 | 3.3 | hypothetical prote   | 334 | 2      | E98118  | 334 | 334    | 334 | 6 | 3.3 | 332 | 2 | AD4730  | hypothetical prote  |
| 262 | 6 | 3.3 | mucin 5AC (clone J   | 335 | 2      | S93163  | 335 | 335    | 335 | 6 | 3.3 | 332 | 2 | T27857  | hypothetical prote  |
| 263 | 6 | 3.3 | probable aliphatic   | 336 | 2      | C96027  | 336 | 336    | 336 | 6 | 3.3 | 333 | 2 | A99582  | transcription regu  |
| 264 | 6 | 3.3 | pantoate beta-alan   | 337 | 2      | F90066  | 337 | 337    | 337 | 6 | 3.3 | 333 | 2 | S66110  | hypothetical prote  |
| 265 | 6 | 3.3 | pyridoxal/pyridoxi   | 338 | 2      | B98040  | 338 | 338    | 338 | 6 | 3.3 | 333 | 2 | T07086  | hypothetical prote  |
| 266 | 6 | 3.3 | pyridoxal/pyridoxi   | 339 | 2      | E85884  | 339 | 339    | 339 | 6 | 3.3 | 334 | 2 | D70918  | hypothetical prote  |
| 267 | 6 | 3.3 | pyridoxal/pyridoxi   | 340 | 2      | A65016  | 340 | 340    | 340 | 6 | 3.3 | 334 | 2 | G02409  | protein kinase C-b  |
| 268 | 6 | 3.3 | paratose kinase      | 341 | 2      | AI0377  | 341 | 341    | 341 | 6 | 3.3 | 334 | 2 | H84730  | phosphate regulato  |
| 269 | 6 | 3.3 | shikimate 5-dehydr   | 342 | 2      | E86844  | 342 | 342    | 342 | 6 | 3.3 | 335 | 2 | A72386  | conserved hypothet  |
| 270 | 6 | 3.3 | hypothetical prote   | 343 | 2      | K1264   | 343 | 343    | 343 | 6 | 3.3 | 335 | 2 | A13111  | hypothetical prote  |
| 271 | 6 | 3.3 | hypothetical prote   | 344 | 2      | A72243  | 344 | 344    | 344 | 6 | 3.3 | 335 | 2 | D98175  | hypothetical prote  |
| 272 | 6 | 3.3 | hypothetical prote   | 345 | 2      | H64538  | 345 | 345    | 345 | 6 | 3.3 | 337 | 2 | T435862 | probable secreted   |
| 273 | 6 | 3.3 | hypothetical prote   | 346 | 2      | F71967  | 346 | 346    | 346 | 6 | 3.3 | 339 | 2 | C42276  | O-methyltransferas  |
| 274 | 6 | 3.3 | hypothetical prote   | 347 | 2      | S77309  | 347 | 347    | 347 | 6 | 3.3 | 342 | 2 | E87605  | conserved hypothet  |
| 275 | 6 | 3.3 | hypothetical prote   | 348 | 2      | G75625  | 348 | 348    | 348 | 6 | 3.3 | 344 | 2 | S41707  | PopAI protein - Ps  |
| 276 | 6 | 3.3 | cell division prot   | 349 | 2      | C75494  | 349 | 349    | 349 | 6 | 3.3 | 344 | 2 | T26228  | hypothetical prote  |
| 277 | 6 | 3.3 | hypothetical prote   | 350 | 2      | F84295  | 350 | 350    | 350 | 6 | 3.3 | 347 | 2 | AD2201  | hypothetical prote  |
| 278 | 6 | 3.3 | hypothetical prote   | 351 | 2      | HS0659  | 351 | 351    | 351 | 6 | 3.3 | 349 | 2 | F95263  | selenide water di   |
| 279 | 6 | 3.3 | homeotic protein H   | 352 | 2      | AH0060  | 352 | 352    | 352 | 6 | 3.3 | 351 | 2 | D84788  | hypothetical prote  |
| 280 | 6 | 3.3 | bis(5'-nucleosyl)-   | 353 | 2      | AC1202  | 353 | 353    | 353 | 6 | 3.3 | 352 | 2 | A39403  | probable homocitra  |
| 281 | 6 | 3.3 | B. subtilis YitL p   | 354 | 2      | AI1559  | 354 | 354    | 354 | 6 | 3.3 | 353 | 2 | T35678  | hypothetical prote  |
| 282 | 6 | 3.3 | B. subtilis YitL p   | 355 | 2      | C72762  | 355 | 355    | 355 | 6 | 3.3 | 354 | 2 | G69171  | hypothetical prote  |
| 283 | 6 | 3.3 | hypothetical prote   | 356 | 2      | B97193  | 356 | 356    | 356 | 6 | 3.3 | 354 | 2 | G69171  | hypothetical prote  |
| 284 | 6 | 3.3 | dihydrodipicolinat   | 357 | 2      | A89791  | 357 | 357    | 357 | 6 | 3.3 | 356 | 1 | WJHU2H  | homeotic protein H  |
| 285 | 6 | 3.3 | hypothetical prote   | 358 | 2      | AB3639  | 358 | 358    | 358 | 6 | 3.3 | 357 | 2 | T02785  | probable homeotic   |
| 286 | 6 | 3.3 | serine proteinase    | 359 | 2      | G87366  | 359 | 359    | 359 | 6 | 3.3 | 357 | 2 | AE2218  | hypothetical prote  |
| 287 | 6 | 3.3 | enoyl-CoA hydratase  | 360 | 2      | F70347  | 360 | 360    | 360 | 6 | 3.3 | 358 | 2 | AE1702  | phospholipase D fa  |
| 288 | 6 | 3.3 | UDP-N-acetylenolpy   | 361 | 2      | D82120  | 361 | 361    | 361 | 6 | 3.3 | 358 | 2 | D72637  | hypothetical prote  |
| 289 | 6 | 3.3 | zinc ABC transport   | 362 | 2      | F70330  | 362 | 362    | 362 | 6 | 3.3 | 359 | 2 | A87023  | hypothetical prote  |
| 290 | 6 | 3.3 | probable arsenical   | 363 | 2      | AB3255  | 363 | 363    | 363 | 6 | 3.3 | 359 | 2 | H70579  | probable murX prot  |
| 291 | 6 | 3.3 | 3-oxoacyl-(acyl-ca   | 364 | 2      | AC82719 | 364 | 364    | 364 | 6 | 3.3 | 359 | 2 | AF0036  | conserved hypothet  |
| 292 | 6 | 3.3 | transcription regu   | 365 | 2      | AB32116 | 365 | 365    | 365 | 6 | 3.3 | 359 | 2 | T44816  | brp protein [impor  |
| 293 | 6 | 3.3 | dihydrodipicolinat   | 366 | 2      | B81696  | 366 | 366    | 366 | 6 | 3.3 | 360 | 2 | F96618  | RNA binding protei  |
| 294 | 6 | 3.3 | 4-hydroxybenzoate    | 367 | 2      | AB6951  | 367 | 367    | 367 | 6 | 3.3 | 361 | 2 | G72636  | hypothetical prote  |
| 295 | 6 | 3.3 | probable periplasm   | 368 | 2      | AB1709  | 368 | 368    | 368 | 6 | 3.3 | 361 | 2 | A72327  | conserved hypothet  |
| 296 | 6 | 3.3 | hypothetical prote   | 369 | 2      | G98035  | 369 | 369    | 369 | 6 | 3.3 | 363 | 2 | S76156  | hypothetical prote  |
| 297 | 6 | 3.3 | thioredoxin-disulf   | 370 | 2      | H70675  | 370 | 370    | 370 | 6 | 3.3 | 363 | 2 | T24596  | hypothetical prote  |
| 298 | 6 | 3.3 | acetaldehyde dehyd   | 371 | 2      | F97639  | 371 | 371    | 371 | 6 | 3.3 | 364 | 2 | B42832  | factor VIII intron  |
| 299 | 6 | 3.3 | hypothetical prote   | 372 | 2      | AG2862  | 372 | 372    | 372 | 6 | 3.3 | 365 | 2 | F84300  | bacteriorhodopsin   |
| 300 | 6 | 3.3 | conserved hypothet   | 373 | 2      | T16457  | 373 | 373    | 373 | 6 | 3.3 | 368 | 2 | F84300  | probable spore ger  |
| 301 | 6 | 3.3 | protein-tyrosine k   | 374 | 2      | TVF5FS  | 374 | 374    | 374 | 6 | 3.3 | 368 | 2 | A96979  | hypothetical glyci  |
| 302 | 6 | 3.3 | pyrroline-5-carbox   | 375 | 2      | S57863  | 375 | 375    | 375 | 6 | 3.3 | 370 | 2 | G70872  | hypothetical prote  |
| 303 | 6 | 3.3 | sensor histidine kin | 376 | 2      | C82425  | 376 | 376    | 376 | 6 | 3.3 | 372 | 2 | T01600  | hypothetical prote  |
| 304 | 6 | 3.3 | probable oxidoredu   | 377 | 2      | F83170  | 377 | 377    | 377 | 6 | 3.3 | 372 | 2 | T01600  | hypothetical prote  |
| 305 | 6 | 3.3 | hypothetical prote   | 378 | 2      | T26043  | 378 | 378    | 378 | 6 | 3.3 | 373 | 2 | B83741  | hypothetical prote  |
| 306 | 6 | 3.3 | activator of 2-hyd   | 379 | 2      | C97199  | 379 | 379    | 379 | 6 | 3.3 | 373 | 2 | B83741  | probable 4-carboxy  |
| 307 | 6 | 3.3 | L-serine ammonia-l   | 380 | 2      | E83310  | 380 | 380    | 380 | 6 | 3.3 | 374 | 2 | T01600  | limulus clotting e  |
| 308 | 6 | 3.3 | probable dihydrosi   | 381 | 2      | D90413  | 381 | 381    | 381 | 6 | 3.3 | 375 | 2 | AG3052  | conserved hypothet  |
| 309 | 6 | 3.3 | thioredoxin reduct   | 382 | 2      | F95390  | 382 | 382    | 382 | 6 | 3.3 | 377 | 2 | T04585  | hypothetical prote  |
| 310 | 6 | 3.3 | probable threonine   | 383 | 2      | T30437  | 383 | 383    | 383 | 6 | 3.3 | 377 | 2 | AG1053  | probable exported   |
| 311 | 6 | 3.3 | hypothetical struc   | 384 | 2      | B95359  | 384 | 384    | 384 | 6 | 3.3 | 379 | 2 | AG1053  | hypothetical prote  |
| 312 | 6 | 3.3 | probable ABC trans   | 385 | 2      | B96606  | 385 | 385    | 385 | 6 | 3.3 | 380 | 2 | B42459  | hypothetical prote  |
| 313 | 6 | 3.3 | conserved hypothet   | 386 | 2      | D95845  | 386 | 386    | 386 | 6 | 3.3 | 381 | 2 | S16506  | hypothetical prote  |
| 314 | 6 | 3.3 | conserved hypothet   | 387 | 2      | F63784  | 387 | 387    | 387 | 6 | 3.3 | 381 | 2 | T52461  | RNA binding protei  |
| 315 | 6 | 3.3 | threonine ammonia-   | 388 | 2      | S38061  | 388 | 388    | 388 | 6 | 3.3 | 383 | 2 | T51466  | hypothetical prote  |
| 316 | 6 | 3.3 | geranylgeranyl dip   | 389 | 2      | F69535  | 389 | 389    | 389 | 6 | 3.3 | 385 | 2 | S53052  | rep8 protein - Lac  |
| 317 | 6 | 3.3 | psp operon transcr   | 390 | 2      | AE0658  | 390 | 390    | 390 | 6 | 3.3 | 385 | 2 | H69154  | hypothetical prote  |
| 318 | 6 | 3.3 | 2'-hydroxydihydrod   | 391 | 2      | T07104  | 391 | 391    | 391 | 6 | 3.3 | 386 | 2 | T30320  | replication protei  |
| 319 | 6 | 3.3 | conserved hypothet   | 392 | 2      | A4613   | 392 | 392    | 392 | 6 | 3.3 | 386 | 2 | T18890  | hypothetical prote  |
| 320 | 6 | 3.3 | hypothetical prote   | 393 | 2      | AB2187  | 393 | 393    | 393 | 6 | 3.3 | 387 | 2 | T29966  | hypothetical prote  |
| 321 | 6 | 3.3 | hypothetical prote   | 394 | 2      | AB2187  | 394 | 394    | 394 | 6 | 3.3 | 389 | 2 | S27200  | proline-rich prote  |
|     |   |     |                      |     |        |         |     |        |     |   |     | 392 | 2 | B69852  | probable NADH2 deh  |
|     |   |     |                      |     |        |         |     |        |     |   |     | 392 | 1 | T51772  | acetyl-CoA C-acety  |

|     |   |     |     |   |        |                    |     |   |     |     |   |        |                     |
|-----|---|-----|-----|---|--------|--------------------|-----|---|-----|-----|---|--------|---------------------|
| 395 | 6 | 3.3 | 396 | 2 | D82584 | penicillin binding | 468 | 6 | 3.3 | 463 | 1 | S74845 | tldd homolog slr08  |
| 396 | 6 | 3.3 | 397 | 1 | S51594 | cytochrome P450 my | 469 | 6 | 3.3 | 463 | 2 | B84279 | hypothetical prote  |
| 397 | 6 | 3.3 | 397 | 2 | T06189 | serpin - barley    | 470 | 6 | 3.3 | 464 | 2 | T33249 | hypothetical prote  |
| 398 | 6 | 3.3 | 400 | 2 | S05945 | granulocyte-macrop | 471 | 6 | 3.3 | 465 | 2 | B85358 | SERINE CARBOXYPEPT  |
| 399 | 6 | 3.3 | 401 | 2 | A13312 | ATP synthase beta  | 472 | 6 | 3.3 | 465 | 2 | H84058 | Xaa-His dipeptidas  |
| 400 | 6 | 3.3 | 401 | 2 | B87684 | hypothetical prote | 473 | 6 | 3.3 | 466 | 2 | AF2274 | hypothetical prote  |
| 401 | 6 | 3.3 | 401 | 2 | C89777 | capsular polysacch | 474 | 6 | 3.3 | 468 | 1 | P2WL51 | L2 protein - human  |
| 402 | 6 | 3.3 | 403 | 2 | D75333 | conserved hypothet | 475 | 6 | 3.3 | 468 | 2 | H87044 | probable solute-bi  |
| 403 | 6 | 3.3 | 404 | 2 | S63352 | probable membrane  | 476 | 6 | 3.3 | 469 | 2 | T45201 | hypothetical prote  |
| 404 | 6 | 3.3 | 405 | 2 | E70545 | hypothetical prote | 477 | 6 | 3.3 | 470 | 2 | D69394 | phosphoribosylamin  |
| 405 | 6 | 3.3 | 405 | 2 | T23321 | hypothetical prote | 478 | 6 | 3.3 | 470 | 2 | T43675 | cog-2 protein - Ga  |
| 406 | 6 | 3.3 | 408 | 2 | A29314 | aspartate kinase ( | 479 | 6 | 3.3 | 472 | 2 | S36548 | L2 protein - human  |
| 407 | 6 | 3.3 | 408 | 2 | D70738 | probable alr prote | 480 | 6 | 3.3 | 474 | 2 | B38634 | tumor necrosis fac  |
| 408 | 6 | 3.3 | 409 | 2 | AF2253 | hypothetical prote | 481 | 6 | 3.3 | 475 | 2 | A12398 | dihydrolipoamide d  |
| 409 | 6 | 3.3 | 410 | 2 | C81185 | Na(+)-translocati  | 482 | 6 | 3.3 | 477 | 2 | A35843 | lipopolysaccharide  |
| 410 | 6 | 3.3 | 410 | 2 | H81918 | probable sodium-tr | 483 | 6 | 3.3 | 478 | 2 | A45796 | dihydrolipoamide d  |
| 411 | 6 | 3.3 | 410 | 2 | F88978 | protein F37B4.4 [i | 484 | 6 | 3.3 | 478 | 2 | A83449 | dihydrolipoamide d  |
| 412 | 6 | 3.3 | 411 | 2 | A48946 | aspartate kinase ( | 485 | 6 | 3.3 | 478 | 2 | T16170 | hypothetical prote  |
| 413 | 6 | 3.3 | 411 | 2 | A72506 | probable tRNA spli | 486 | 6 | 3.3 | 481 | 2 | A54136 | lipopolysaccharide  |
| 414 | 6 | 3.3 | 413 | 2 | AF0393 | NADH2 dehydrogen   | 487 | 6 | 3.3 | 484 | 2 | AF3107 | xylokinase limpo    |
| 415 | 6 | 3.3 | 416 | 2 | D81332 | phosphoribosylamin | 488 | 6 | 3.3 | 485 | 2 | T24677 | hypothetical prote  |
| 416 | 6 | 3.3 | 417 | 2 | A49129 | neuroblast cell li | 489 | 6 | 3.3 | 487 | 2 | T49080 | serine-type carbox  |
| 417 | 6 | 3.3 | 419 | 1 | S75867 | phosphoribosylamin | 490 | 6 | 3.3 | 487 | 2 | H82298 | D-alanyl-D-alanine  |
| 418 | 6 | 3.3 | 420 | 2 | G95005 | phosphoribosylamin | 491 | 6 | 3.3 | 488 | 2 | F98179 | homoeitic protein H |
| 419 | 6 | 3.3 | 420 | 2 | D97878 | phosphoribosylamin | 492 | 6 | 3.3 | 488 | 2 | A55180 | hypothetical prote  |
| 420 | 6 | 3.3 | 424 | 2 | E75362 | glutamate dehydrog | 493 | 6 | 3.3 | 490 | 2 | T36920 | conserved hypothet  |
| 421 | 6 | 3.3 | 424 | 2 | B70365 | phosphoribosylamin | 494 | 6 | 3.3 | 490 | 2 | F87443 | cosmid ZK353 prote  |
| 422 | 6 | 3.3 | 425 | 2 | AG2244 | phosphoribosylamin | 495 | 6 | 3.3 | 491 | 2 | S44657 | succinylglutamic s  |
| 423 | 6 | 3.3 | 426 | 2 | T36948 | Ig epsilon-chain - | 496 | 6 | 3.3 | 492 | 2 | B64934 | sodium/proline sym  |
| 424 | 6 | 3.3 | 426 | 2 | B69876 | acetylornithine de | 497 | 6 | 3.3 | 492 | 2 | H69670 | amino acid transpo  |
| 425 | 6 | 3.3 | 426 | 2 | S51016 | sodium-translocati | 498 | 6 | 3.3 | 493 | 2 | S52421 | hypothetical prote  |
| 426 | 6 | 3.3 | 426 | 2 | G96528 | protein F27J15.2 [ | 499 | 6 | 3.3 | 493 | 2 | C87362 | protein T08H10.1 [  |
| 427 | 6 | 3.3 | 427 | 2 | A13441 | phosphoribosylamin | 500 | 6 | 3.3 | 496 | 2 | B89027 | N-acetylglucosamin  |
| 428 | 6 | 3.3 | 427 | 2 | F82094 | sodium-translocati | 501 | 6 | 3.3 | 496 | 2 | A54770 | probable bacteriop  |
| 429 | 6 | 3.3 | 427 | 2 | G82827 | 6-phosphofructokin | 502 | 6 | 3.3 | 497 | 2 | AD0152 | sensor histidine k  |
| 430 | 6 | 3.3 | 428 | 1 | EH0709 | Ig epsilon chain C | 503 | 6 | 3.3 | 497 | 2 | G82447 | xylose kinase xylB  |
| 431 | 6 | 3.3 | 428 | 2 | AH0709 | succinylglutamic s | 504 | 6 | 3.3 | 498 | 2 | D83994 | hypothetical prote  |
| 432 | 6 | 3.3 | 429 | 1 | C64100 | phosphoribosylamin | 505 | 6 | 3.3 | 500 | 2 | S61591 | hypothetical prote  |
| 433 | 6 | 3.3 | 429 | 2 | B82344 | phosphoribosylamin | 506 | 6 | 3.3 | 501 | 2 | A25338 | nicotinic acetylch  |
| 434 | 6 | 3.3 | 430 | 2 | S36233 | leucoanthocyanidin | 507 | 6 | 3.3 | 501 | 2 | S13873 | nicotinic acetylch  |
| 435 | 6 | 3.3 | 430 | 2 | T45151 | dihydroorotase (EC | 508 | 6 | 3.3 | 501 | 2 | T36051 | probable cytochrom  |
| 436 | 6 | 3.3 | 430 | 2 | A89469 | protein F07G6.3 [i | 509 | 6 | 3.3 | 501 | 2 | F87550 | hypothetical prote  |
| 437 | 6 | 3.3 | 431 | 2 | E81053 | seryl-tRNA synthet | 510 | 6 | 3.3 | 503 | 2 | A40583 | heat shock transcr  |
| 438 | 6 | 3.3 | 431 | 2 | E81822 | serine-tRNA ligase | 511 | 6 | 3.3 | 505 | 2 | T41303 | hypothetical prote  |
| 439 | 6 | 3.3 | 434 | 1 | A45027 | adenylosuccinate s | 512 | 6 | 3.3 | 507 | 2 | T18639 | hypothetical prote  |
| 440 | 6 | 3.3 | 434 | 2 | T37670 | adenylosuccinate s | 513 | 6 | 3.3 | 509 | 2 | A47174 | serotonin receptor  |
| 441 | 6 | 3.3 | 437 | 2 | AG2945 | hypothetical prote | 514 | 6 | 3.3 | 512 | 2 | AF0977 | lactaldehyde dehyd  |
| 442 | 6 | 3.3 | 438 | 2 | T35355 | hypothetical prote | 515 | 6 | 3.3 | 515 | 2 | AD1048 | conserved hypothet  |
| 443 | 6 | 3.3 | 439 | 2 | D82244 | sensory box sensor | 516 | 6 | 3.3 | 515 | 2 | G91271 | hypothetical prote  |
| 444 | 6 | 3.3 | 441 | 2 | E75357 | probable periplasm | 517 | 6 | 3.3 | 515 | 2 | S56392 | hypothetical 54.7K  |
| 445 | 6 | 3.3 | 441 | 2 | AB1542 | PTS system, Lichen | 518 | 6 | 3.3 | 515 | 2 | G86112 | hypothetical prote  |
| 446 | 6 | 3.3 | 441 | 2 | AD1184 | PTS system, Lichen | 519 | 6 | 3.3 | 515 | 2 | AC2635 | glycosyltransferas  |
| 447 | 6 | 3.3 | 445 | 2 | D83053 | phosphoglucosamine | 520 | 6 | 3.3 | 516 | 2 | A84081 | hypothetical prote  |
| 448 | 6 | 3.3 | 445 | 2 | S06992 | wall-associated pr | 521 | 6 | 3.3 | 518 | 2 | F66814 | bifunctional purin  |
| 449 | 6 | 3.3 | 446 | 2 | B87469 | major facilitator  | 522 | 6 | 3.3 | 518 | 2 | D64244 | H+-transporting tw  |
| 450 | 6 | 3.3 | 447 | 2 | G84687 | probable disease r | 523 | 6 | 3.3 | 518 | 2 | S62847 | H+-transporting tw  |
| 451 | 6 | 3.3 | 448 | 2 | AC3602 | transporter, mfs s | 524 | 6 | 3.3 | 519 | 2 | T27880 | hypothetical prote  |
| 452 | 6 | 3.3 | 448 | 2 | S62679 | phosphomannomutase | 525 | 6 | 3.3 | 522 | 2 | T28323 | ORF MSV162 probabl  |
| 453 | 6 | 3.3 | 448 | 2 | AB2679 | nitric-oxide reduc | 526 | 6 | 3.3 | 524 | 2 | AG3328 | proteinase do (EC   |
| 454 | 6 | 3.3 | 449 | 2 | AE3634 | phosphoglucumutase | 527 | 6 | 3.3 | 525 | 2 | T10478 | probable 1,2-diacy  |
| 455 | 6 | 3.3 | 450 | 2 | AF1339 | phosphoglucumutase | 528 | 6 | 3.3 | 526 | 1 | A34164 | cholesterol monoox  |
| 456 | 6 | 3.3 | 450 | 2 | AD1710 | periplasmic sorbit | 529 | 6 | 3.3 | 526 | 2 | B70859 | hypothetical prote  |
| 457 | 6 | 3.3 | 450 | 2 | C98337 | probable cyclin [i | 530 | 6 | 3.3 | 529 | 2 | A41137 | heat shock transcr  |
| 458 | 6 | 3.3 | 452 | 2 | F86289 | hypothetical prote | 531 | 6 | 3.3 | 529 | 2 | A90328 | hypothetical prote  |
| 459 | 6 | 3.3 | 452 | 2 | E72387 | hypothetical prote | 532 | 6 | 3.3 | 530 | 1 | T44892 | probable dihydroli  |
| 460 | 6 | 3.3 | 453 | 2 | T30239 | hypothetical prote | 533 | 6 | 3.3 | 530 | 2 | QC7983 | peptidoglycan reco  |
| 461 | 6 | 3.3 | 454 | 2 | A69017 | dihydroorotase - M | 534 | 6 | 3.3 | 532 | 1 | UF7983 | neurofilament trip  |
| 462 | 6 | 3.3 | 454 | 2 | T39297 | probable atpase -  | 535 | 6 | 3.3 | 532 | 2 | AF7064 | D-3-phosphoglycera  |
| 463 | 6 | 3.3 | 459 | 2 | I48854 | gene murine tumour | 536 | 6 | 3.3 | 533 | 2 | T00742 | ubiquitin-binding   |
| 464 | 6 | 3.3 | 459 | 2 | T35317 | probable serine/th | 537 | 6 | 3.3 | 533 | 2 | B96642 | hypothetical prote  |
| 465 | 6 | 3.3 | 459 | 2 | T45911 | hypothetical prote | 538 | 6 | 3.3 | 534 | 2 | A44097 | methylnalonate-sem  |
| 466 | 6 | 3.3 | 461 | 1 | A35356 | tumor necrosis fac | 539 | 6 | 3.3 | 535 | 1 | A46600 | hypothetical prote  |
| 467 | 6 | 3.3 | 461 | 2 | AG0452 | aspartate kinase ( | 540 | 6 | 3.3 | 537 | 2 | T02982 | probable sucrose t  |

|     |   |     |     |   |        |                     |     |   |     |     |   |        |                    |
|-----|---|-----|-----|---|--------|---------------------|-----|---|-----|-----|---|--------|--------------------|
| 541 | 6 | 3.3 | 541 | 2 | AF2483 | hypothetical prote  | 614 | 6 | 3.3 | 670 | 2 | AE2111 | general secretion  |
| 542 | 6 | 3.3 | 542 | 1 | S47809 | aldehyde dehydroge  | 615 | 6 | 3.3 | 672 | 2 | A56765 | sodium-glucose cot |
| 543 | 6 | 3.3 | 543 | 2 | G86033 | aldehyde dehydroge  | 616 | 6 | 3.3 | 672 | 2 | A42251 | nucleoside transpo |
| 544 | 6 | 3.3 | 544 | 2 | H91186 | aldehyde dehydroge  | 617 | 6 | 3.3 | 672 | 2 | S74374 | general secretion  |
| 545 | 6 | 3.3 | 545 | 2 | T30491 | hypothetical prote  | 618 | 6 | 3.3 | 674 | 2 | E82954 | conserved hypothet |
| 546 | 6 | 3.3 | 546 | 2 | T01100 | reverse transcript  | 619 | 6 | 3.3 | 675 | 2 | E81101 | transporter, BCCT  |
| 547 | 6 | 3.3 | 547 | 2 | T16145 | hypothetical prote  | 620 | 6 | 3.3 | 677 | 2 | F64139 | guanosine-3',5'-bi |
| 548 | 6 | 3.3 | 548 | 2 | T45535 | agag protein [impo  | 621 | 6 | 3.3 | 678 | 2 | T50256 | probable vacuolar  |
| 549 | 6 | 3.3 | 549 | 2 | S64725 | probable lipoprote  | 622 | 6 | 3.3 | 678 | 2 | B83154 | conserved hypothet |
| 550 | 6 | 3.3 | 550 | 2 | H70977 | hypothetical prote  | 623 | 6 | 3.3 | 680 | 2 | H90558 | hypothetical prote |
| 551 | 6 | 3.3 | 551 | 2 | S63177 | mannosyl transfera  | 624 | 6 | 3.3 | 684 | 1 | RNLVC1 | DNA-directed RNA p |
| 552 | 6 | 3.3 | 552 | 2 | T47506 | hypothetical prote  | 625 | 6 | 3.3 | 689 | 2 | AC1927 | hypothetical prote |
| 553 | 6 | 3.3 | 553 | 2 | G84221 | NADH dehydrogenase  | 626 | 6 | 3.3 | 690 | 2 | T46090 | hypothetical prote |
| 554 | 6 | 3.3 | 554 | 2 | S78224 | virulence-associat  | 627 | 6 | 3.3 | 691 | 2 | B75622 | hypothetical prote |
| 555 | 6 | 3.3 | 555 | 2 | S57124 | Cmp synthase (EC 6  | 628 | 6 | 3.3 | 694 | 2 | G01161 | hypothetical prote |
| 556 | 6 | 3.3 | 556 | 2 | T49197 | hypothetical prote  | 629 | 6 | 3.3 | 700 | 2 | T40596 | glutamate-ammonia  |
| 557 | 6 | 3.3 | 557 | 2 | AF3639 | pseudouridylyate sy | 630 | 6 | 3.3 | 700 | 2 | D75508 | serine/threonine p |
| 558 | 6 | 3.3 | 558 | 2 | A55377 | CPE-binding protei  | 631 | 6 | 3.3 | 701 | 2 | H82978 | guanosine-3',5'-bi |
| 559 | 6 | 3.3 | 559 | 2 | F75511 | probable 5'-nucleo  | 632 | 6 | 3.3 | 705 | 2 | T31157 | hypothetical prote |
| 560 | 6 | 3.3 | 560 | 2 | E64427 | hypothetical prote  | 633 | 6 | 3.3 | 706 | 2 | AH0670 | probable TonB-depe |
| 561 | 6 | 3.3 | 561 | 2 | D72597 | hypothetical prote  | 634 | 6 | 3.3 | 706 | 2 | S43581 | C28A5.6 protein (C |
| 562 | 6 | 3.3 | 562 | 2 | T08930 | hypothetical prote  | 635 | 6 | 3.3 | 707 | 2 | A46302 | PTB-associated spl |
| 563 | 6 | 3.3 | 563 | 2 | AD1209 | N-acetylmuramoyl-L  | 636 | 6 | 3.3 | 712 | 1 | YSHUT  | threonine-tRNA lig |
| 564 | 6 | 3.3 | 564 | 2 | S34340 | NPL4 protein - yea  | 637 | 6 | 3.3 | 712 | 2 | T47243 | amino-acid N-acety |
| 565 | 6 | 3.3 | 565 | 2 | E69322 | dolichol-P-glucose  | 638 | 6 | 3.3 | 712 | 2 | T48961 | hypothetical prote |
| 566 | 6 | 3.3 | 566 | 2 | T49206 | phosphoinositide-s  | 639 | 6 | 3.3 | 715 | 2 | B75135 | DNA helicase relat |
| 567 | 6 | 3.3 | 567 | 2 | A90394 | bps2 protein homol  | 640 | 6 | 3.3 | 715 | 2 | JC4908 | alkaline serine pr |
| 568 | 6 | 3.3 | 568 | 2 | PW0007 | chaperonin 62.5K b  | 641 | 6 | 3.3 | 718 | 2 | T02531 | probable SCARECROW |
| 569 | 6 | 3.3 | 569 | 2 | S09498 | virulence-associat  | 642 | 6 | 3.3 | 719 | 2 | JC1200 | alpha-glucosidase  |
| 570 | 6 | 3.3 | 570 | 2 | S22664 | virulence-associat  | 643 | 6 | 3.3 | 719 | 2 | I39814 | insecticidal prote |
| 571 | 6 | 3.3 | 571 | 2 | T07616 | probable beta-fruc  | 644 | 6 | 3.3 | 719 | 2 | S25383 | parasporal crystal |
| 572 | 6 | 3.3 | 572 | 2 | S15215 | virulence-associat  | 645 | 6 | 3.3 | 719 | 2 | I39815 | insecticidal prote |
| 573 | 6 | 3.3 | 573 | 2 | T06412 | probable chaperoni  | 646 | 6 | 3.3 | 719 | 2 | AC0867 | 2-acylglycerophosp |
| 574 | 6 | 3.3 | 574 | 2 | B41627 | furin (EC 3.4.21.7  | 647 | 6 | 3.3 | 720 | 2 | G64230 | stringent response |
| 575 | 6 | 3.3 | 575 | 2 | A45195 | adenylcyclase cy    | 648 | 6 | 3.3 | 724 | 2 | H69780 | antibiotic transpo |
| 576 | 6 | 3.3 | 576 | 2 | S00962 | hypothetical prote  | 649 | 6 | 3.3 | 725 | 2 | T10464 | fatty acid oxidati |
| 577 | 6 | 3.3 | 577 | 2 | H64496 | hypothetical prote  | 650 | 6 | 3.3 | 727 | 2 | H69724 | DNA topoisomerase  |
| 578 | 6 | 3.3 | 578 | 2 | T07733 | probable chaperoni  | 651 | 6 | 3.3 | 729 | 2 | S35532 | hnRNA-binding prot |
| 579 | 6 | 3.3 | 579 | 2 | JT0901 | chaperonin 60 beta  | 652 | 6 | 3.3 | 732 | 2 | AD0014 | primosomal protein |
| 580 | 6 | 3.3 | 580 | 2 | B96597 | Rubisco subunit bi  | 653 | 6 | 3.3 | 733 | 2 | S73767 | stringent response |
| 581 | 6 | 3.3 | 581 | 2 | AE2437 | prolyl-tRNA synthe  | 654 | 6 | 3.3 | 734 | 2 | B90599 | atp-dependent heli |
| 582 | 6 | 3.3 | 582 | 2 | S57552 | hypothetical prote  | 655 | 6 | 3.3 | 735 | 2 | S61238 | hypothetical prote |
| 583 | 6 | 3.3 | 583 | 2 | JQ1462 | phosphoenolpyruvat  | 656 | 6 | 3.3 | 736 | 2 | D86271 | protein F16A14.2 [ |
| 584 | 6 | 3.3 | 584 | 2 | E81451 | 1-deoxyxylulose-5-  | 657 | 6 | 3.3 | 742 | 2 | T25415 | hypothetical prote |
| 585 | 6 | 3.3 | 585 | 2 | JX0140 | cytochrome-c oxida  | 658 | 6 | 3.3 | 743 | 1 | G02270 | alpha-N-acetylgluc |
| 586 | 6 | 3.3 | 586 | 2 | S75447 | proline-tRNA ligas  | 659 | 6 | 3.3 | 746 | 2 | AD1622 | probable integral  |
| 587 | 6 | 3.3 | 587 | 2 | T20450 | hypothetical prote  | 660 | 6 | 3.3 | 748 | 2 | G83927 | alpha-galactosidas |
| 588 | 6 | 3.3 | 588 | 2 | S56214 | probable membrane   | 661 | 6 | 3.3 | 750 | 2 | S77636 | exopolysaccharide  |
| 589 | 6 | 3.3 | 589 | 2 | AH0521 | dihydrolipoamide a  | 662 | 6 | 3.3 | 754 | 2 | S62512 | probable cysteinyl |
| 590 | 6 | 3.3 | 590 | 2 | G90643 | dihydrolipoamide S  | 663 | 6 | 3.3 | 757 | 2 | T34362 | hypothetical prote |
| 591 | 6 | 3.3 | 591 | 2 | G90643 | hypothetical prote  | 664 | 6 | 3.3 | 761 | 2 | T24230 | hypothetical prote |
| 592 | 6 | 3.3 | 592 | 2 | G85494 | hypothetical prote  | 665 | 6 | 3.3 | 761 | 2 | AB1848 | hypothetical prote |
| 593 | 6 | 3.3 | 593 | 2 | T04179 | hypothetical prote  | 666 | 6 | 3.3 | 762 | 2 | D87403 | hypothetical prote |
| 594 | 6 | 3.3 | 594 | 2 | G89836 | ABC transporter pe  | 667 | 6 | 3.3 | 765 | 2 | T40674 | protein transport  |
| 595 | 6 | 3.3 | 595 | 2 | C82079 | pyruvate dehydroge  | 668 | 6 | 3.3 | 771 | 2 | T21633 | hypothetical prote |
| 596 | 6 | 3.3 | 596 | 2 | T23874 | hypothetical prote  | 669 | 6 | 3.3 | 772 | 2 | T08226 | hypothetical prote |
| 597 | 6 | 3.3 | 597 | 2 | A82738 | heat shock protein  | 670 | 6 | 3.3 | 775 | 1 | JQ1539 | outer layer protei |
| 598 | 6 | 3.3 | 598 | 2 | T28866 | hypothetical prote  | 671 | 6 | 3.3 | 781 | 1 | TVFFDF | protein kinase Dra |
| 599 | 6 | 3.3 | 599 | 2 | C95991 | conserved hypothet  | 672 | 6 | 3.3 | 783 | 2 | T35389 | probable serine-th |
| 600 | 6 | 3.3 | 600 | 2 | B87089 | probable apolipop   | 673 | 6 | 3.3 | 784 | 2 | H87398 | transketolase fami |
| 601 | 6 | 3.3 | 601 | 2 | ODPPI  | cytochrome-c oxida  | 674 | 6 | 3.3 | 784 | 2 | AC1091 | 5'-nucleotidase, p |
| 602 | 6 | 3.3 | 602 | 2 | S07751 | cytochrome-c oxida  | 675 | 6 | 3.3 | 785 | 2 | AB1455 | 5'-nucleotidase, p |
| 603 | 6 | 3.3 | 603 | 2 | B97417 | probable membrane   | 676 | 6 | 3.3 | 786 | 2 | S71091 | acetyl-CoA carboxy |
| 604 | 6 | 3.3 | 604 | 2 | T38545 | hypothetical prote  | 677 | 6 | 3.3 | 789 | 2 | AE2688 | Na+/H+ antiporter  |
| 605 | 6 | 3.3 | 605 | 2 | T36419 | hypothetical prote  | 678 | 6 | 3.3 | 789 | 2 | H97469 | probable NADH dehy |
| 606 | 6 | 3.3 | 606 | 2 | AE3230 | hydantoin utilizat  | 679 | 6 | 3.3 | 798 | 2 | T34248 | hypothetical prote |
| 607 | 6 | 3.3 | 607 | 2 | G71697 | probable soluble l  | 680 | 6 | 3.3 | 799 | 2 | JH0797 | castor protein - f |
| 608 | 6 | 3.3 | 608 | 2 | E86787 | hypothetical prote  | 681 | 6 | 3.3 | 801 | 2 | B86673 | penicillin-binding |
| 609 | 6 | 3.3 | 609 | 2 | T01301 | RNA-directed DNA p  | 682 | 6 | 3.3 | 802 | 1 | A26343 | phosphoribosylamin |
| 610 | 6 | 3.3 | 610 | 2 | E81839 | probable transmem   | 683 | 6 | 3.3 | 804 | 2 | AG0565 | probable membrane  |
| 611 | 6 | 3.3 | 611 | 2 | S40098 | hypothetical prote  | 684 | 6 | 3.3 | 804 | 2 | G64780 | probable membrane  |
| 612 | 6 | 3.3 | 612 | 2 | AG1026 | topoisomerase B [i  | 685 | 6 | 3.3 | 804 | 2 | A85549 | probable oxidoredu |
| 613 | 6 | 3.3 | 613 | 2 | A84742 | probable polygalac  | 686 | 6 | 3.3 | 804 | 2 | G90698 | probable oxidoredu |





|     |    |   |        |                    |     |   |     |    |   |        |                    |
|-----|----|---|--------|--------------------|-----|---|-----|----|---|--------|--------------------|
| 833 | 10 | 2 | F60527 | sperm-activating p | 906 | 5 | 2.8 | 62 | 2 | AC1807 | hypothetical prote |
| 834 | 10 | 2 | F60527 | sperm-activating p | 907 | 5 | 2.8 | 63 | 2 | C83524 | probable cold-shoc |
| 835 | 10 | 2 | E60788 | sperm-activating p | 908 | 5 | 2.8 | 63 | 2 | T08130 | oleosin-like prote |
| 836 | 10 | 2 | A60589 | sperm-activating p | 909 | 5 | 2.8 | 63 | 2 | AC2145 | hypothetical prote |
| 837 | 11 | 2 | I52304 | gene rSSR4 protei  | 910 | 5 | 2.8 | 64 | 2 | C48555 | hypothetical prote |
| 838 | 14 | 2 | F50249 | porin - rice (stra | 911 | 5 | 2.8 | 64 | 2 | B42456 | hypothetical prote |
| 839 | 15 | 2 | G41299 | T-cell receptor al | 912 | 5 | 2.8 | 64 | 2 | AI2881 | hypothetical prote |
| 840 | 15 | 2 | A36279 | chemoattractant pr | 913 | 5 | 2.8 | 65 | 2 | T06349 | glucan endo-1,3-be |
| 841 | 17 | 2 | A60889 | olfactory glycopro | 914 | 5 | 2.8 | 65 | 2 | AH1024 | hypothetical prote |
| 842 | 18 | 2 | S46418 | NTL1 protein - cur | 915 | 5 | 2.8 | 65 | 2 | S74783 | hypothetical prote |
| 843 | 20 | 2 | A49048 | T-cell receptor be | 916 | 5 | 2.8 | 65 | 2 | T33300 | hypothetical prote |
| 844 | 24 | 2 | PH1909 | T-cell receptor al | 917 | 5 | 2.8 | 65 | 2 | F97800 | hypothetical prote |
| 845 | 27 | 2 | S00347 | triacylglycerol li | 918 | 5 | 2.8 | 66 | 2 | H81813 | hypothetical prote |
| 846 | 30 | 2 | F45095 | photosystem I ligh | 919 | 5 | 2.8 | 66 | 2 | T04546 | hypothetical prote |
| 847 | 30 | 2 | B95020 | hypothetical prote | 920 | 5 | 2.8 | 66 | 2 | T17015 | metallothionein-li |
| 848 | 31 | 2 | AB0049 | transposase (parti | 921 | 5 | 2.8 | 66 | 2 | T18859 | hypothetical prote |
| 849 | 32 | 2 | F82833 | hypothetical prote | 922 | 5 | 2.8 | 67 | 2 | I69230 | hypothetical prote |
| 850 | 33 | 2 | B39509 | mannose-specific l | 923 | 5 | 2.8 | 67 | 2 | AB0396 | hypothetical prote |
| 851 | 36 | 2 | A28503 | neuropeptide H - b | 924 | 5 | 2.8 | 68 | 2 | H87331 | hypothetical prote |
| 852 | 36 | 2 | A81164 | hypothetical prote | 925 | 5 | 2.8 | 68 | 2 | H72284 | hypothetical prote |
| 853 | 38 | 2 | A45588 | kunitz trypsin inh | 926 | 5 | 2.8 | 69 | 1 | MEC77  | microcin B17 precu |
| 854 | 40 | 1 | F2R27  | photosystem II pro | 927 | 5 | 2.8 | 69 | 2 | H95016 | hypothetical prote |
| 855 | 40 | 2 | S58565 | probable photosyst | 928 | 5 | 2.8 | 70 | 2 | H83262 | ribosome modulatio |
| 856 | 40 | 2 | T07478 | probable photosyst | 929 | 5 | 2.8 | 70 | 2 | B64497 | hypothetical prote |
| 857 | 42 | 2 | C38578 | protein kinase 3 { | 930 | 5 | 2.8 | 71 | 2 | S51665 | conserved hypothe  |
| 858 | 42 | 2 | G90028 | hypothetical prote | 931 | 5 | 2.8 | 71 | 2 | E96999 | ribosomal protein  |
| 859 | 42 | 2 | AD0269 | hypothetical prote | 932 | 5 | 2.8 | 71 | 2 | E96999 | hydrogenase expres |
| 860 | 45 | 1 | RUDVD  | rubredoxin [valida | 933 | 5 | 2.8 | 71 | 2 | A95248 | transcription regu |
| 861 | 46 | 2 | S16198 | diuretic peptide - | 934 | 5 | 2.8 | 71 | 2 | F83214 | probable molybdopt |
| 862 | 46 | 2 | AE0009 | hypothetical prote | 935 | 5 | 2.8 | 71 | 2 | H81322 | hypothetical prote |
| 863 | 47 | 2 | S10862 | hypothetical prote | 936 | 5 | 2.8 | 71 | 2 | D70638 | hypothetical prote |
| 864 | 47 | 2 | H90945 | hypothetical prote | 937 | 5 | 2.8 | 71 | 2 | H69284 | conserved hypothe  |
| 865 | 47 | 2 | C85794 | hypothetical prote | 938 | 5 | 2.8 | 71 | 2 | F98112 | conserved hypothe  |
| 866 | 47 | 2 | B64944 | hypothetical prote | 939 | 5 | 2.8 | 72 | 2 | T07073 | metallothionein ty |
| 867 | 48 | 2 | Z26612 | spleenin - human   | 940 | 5 | 2.8 | 72 | 2 | T07105 | metallothionein-li |
| 868 | 48 | 2 | A26612 | thymopoietin - hum | 941 | 5 | 2.8 | 72 | 2 | S68883 | light-harvesting p |
| 869 | 49 | 1 | TOB01  | thymopoietin I - b | 942 | 5 | 2.8 | 72 | 2 | AD0561 | haemolysin express |
| 870 | 49 | 1 | TOB02  | thymopoietin II -  | 943 | 5 | 2.8 | 72 | 2 | T16756 | hypothetical prote |
| 871 | 49 | 1 | TOB03  | thymopoietin III - | 944 | 5 | 2.8 | 72 | 2 | T18606 | hypothetical prote |
| 872 | 49 | 2 | S70651 | leukotriene-A4 hyd | 945 | 5 | 2.8 | 72 | 2 | T29015 | hypothetical prote |
| 873 | 49 | 2 | G69338 | conserved hypothe  | 946 | 5 | 2.8 | 73 | 1 | S61830 | subtilisin/chymotr |
| 874 | 50 | 2 | A71570 | hypothetical prote | 947 | 5 | 2.8 | 73 | 2 | A60136 | keratin, scale (cl |
| 875 | 51 | 2 | T26197 | hypothetical prote | 948 | 5 | 2.8 | 73 | 2 | T07109 | metallothionein-li |
| 876 | 52 | 2 | S77766 | probable ornithine | 949 | 5 | 2.8 | 73 | 2 | B64538 | cbb3-type cytochro |
| 877 | 52 | 2 | A26677 | purpurin - chicken | 950 | 5 | 2.8 | 73 | 2 | AB1413 | hypothetical prote |
| 878 | 52 | 2 | G97730 | hypothetical prote | 951 | 5 | 2.8 | 73 | 2 | AI1788 | hypothetical prote |
| 879 | 53 | 2 | QJ1664 | hypothetical 6.2K  | 952 | 5 | 2.8 | 73 | 2 | AI3065 | hypothetical prote |
| 880 | 53 | 2 | H82563 | hypothetical prote | 953 | 5 | 2.8 | 73 | 2 | G98220 | hypothetical prote |
| 881 | 54 | 2 | F00291 | lipoygenase (EC 1  | 954 | 5 | 2.8 | 74 | 2 | T16979 | metallothionein-li |
| 882 | 54 | 2 | T42357 | hypothetical prote | 955 | 5 | 2.8 | 74 | 2 | S60849 | M protein precurs  |
| 883 | 54 | 2 | S35183 | cytochrome p450 (c | 956 | 5 | 2.8 | 74 | 2 | AI0546 | hypothetical prote |
| 884 | 55 | 2 | F75314 | ribosomal protein  | 957 | 5 | 2.8 | 74 | 2 | G84780 | hypothetical prote |
| 885 | 55 | 2 | T45700 | hypothetical prote | 958 | 5 | 2.8 | 74 | 2 | G75443 | hypothetical prote |
| 886 | 55 | 2 | T42641 | syntaxin C - Caeno | 959 | 5 | 2.8 | 74 | 2 | S72758 | hypothetical prote |
| 887 | 55 | 2 | AH1443 | phage protein homo | 960 | 5 | 2.8 | 74 | 2 | E69028 | hypothetical prote |
| 888 | 55 | 2 | B22175 | heat shock protein | 961 | 5 | 2.8 | 74 | 2 | T33086 | hypothetical prote |
| 889 | 56 | 2 | F81824 | hypothetical prote | 962 | 5 | 2.8 | 74 | 2 | G85899 | probable oxidoredu |
| 890 | 56 | 2 | D81053 | hypothetical prote | 963 | 5 | 2.8 | 74 | 2 | H97797 | hypothetical prote |
| 891 | 56 | 2 | PC1011 | PC7 protein - mous | 964 | 5 | 2.8 | 75 | 2 | AB1193 | hypothetical prote |
| 892 | 57 | 2 | AB0756 | hypothetical prote | 965 | 5 | 2.8 | 76 | 2 | S65527 | H+-exporting ATPas |
| 893 | 57 | 2 | JN0739 | hypothetical 6.4K  | 966 | 5 | 2.8 | 76 | 2 | T24627 | hypothetical prote |
| 894 | 57 | 2 | T22546 | hypothetical prote | 967 | 5 | 2.8 | 76 | 2 | T28248 | hypothetical prote |
| 895 | 57 | 2 | S77749 | hypothetical prote | 968 | 5 | 2.8 | 77 | 2 | S57862 | ORF MSV087 probabl |
| 896 | 59 | 2 | S75907 | hypothetical prote | 969 | 5 | 2.8 | 77 | 2 | AP2328 | metallothionein 2b |
| 897 | 59 | 2 | E69950 | hypothetical prote | 970 | 5 | 2.8 | 77 | 2 | D82819 | hypothetical prote |
| 898 | 59 | 2 | T22272 | hypothetical prote | 971 | 5 | 2.8 | 79 | 2 | B41793 | dipeptidyl aminope |
| 899 | 59 | 2 | AF1732 | hypothetical prote | 972 | 5 | 2.8 | 79 | 2 | T03713 | reverse transcript |
| 900 | 60 | 2 | D44088 | homeotic protein H | 973 | 5 | 2.8 | 80 | 2 | B90521 | hypothetical prote |
| 901 | 60 | 2 | S27152 | GPI-anchored epidi | 974 | 5 | 2.8 | 80 | 2 | G96841 | hypothetical prote |
| 902 | 61 | 2 | T06751 | hypothetical prote | 975 | 5 | 2.8 | 81 | 2 | B97253 | foF1-type ATP synt |
| 903 | 61 | 2 | S18766 | GPI-anchored epidi | 976 | 5 | 2.8 | 81 | 2 | JC4147 | intestinal trefoil |
| 904 | 61 | 2 | C82591 | hypothetical prote | 977 | 5 | 2.8 | 81 | 2 | S47426 | envelope protein s |
| 905 | 62 | 2 | T44401 | ribosomal protein  | 978 | 5 | 2.8 | 81 | 2 | C60076 | hypothetical prote |

979 5 2.8 81 2 F96904 hypothetical prote  
 980 5 2.8 81 2 B6271 hypothetical prote  
 981 5 2.8 82 2 S60832 M protein precurs  
 982 5 2.8 82 2 A49276 42k protein - Extr  
 983 5 2.8 82 2 JQ1723 small membrane pro  
 984 5 2.8 82 2 S24280 hypothetical prote  
 985 5 2.8 82 2 C60007 hypothetical prote  
 986 5 2.8 82 2 S01741 hypothetical prote  
 987 5 2.8 82 2 T52378 probable transport  
 988 5 2.8 82 2 A97829 hypothetical prote  
 989 5 2.8 83 1 R3RT21 ribosomal protein  
 990 5 2.8 83 2 I49494 B-1 alpha-amylase  
 991 5 2.8 83 2 JW0086 trichothecene 3-O-  
 992 5 2.8 83 2 S34108 ribosomal protein  
 993 5 2.8 83 2 S05975 tetracenomycin C-p  
 994 5 2.8 83 2 T17726 hypothetical prote  
 995 5 2.8 83 2 B23780 bone protein II -  
 996 5 2.8 84 2 D75414 ribosomal protein  
 997 5 2.8 84 2 F84388 hypothetical prote  
 998 5 2.8 84 2 S33999 EP84R protein - Af  
 999 5 2.8 84 2 AB1295 hypothetical prote  
 1000 5 2.8 84 2 AH1666 hypothetical prote

## ALIGNMENTS

RESULT 1  
 A10409 superoxide dismutase [EC 1.15.1.1] precursor [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 16-Aug-2004  
 C:Accession: A10409  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: A10409  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-201 <KUR>  
 A:Cross-references: UNIPROT:Q8ZBN3; GB:AL590842; PIDN:CAC92605.1; PID:gl15981301; GSPDB:G  
 C:Genetics:  
 A:Gene: sodC  
 C:Superfamily: Superoxide dismutase [Cu-Zn]  
 C:Keywords: oxidoreductase

Query Match 6.1%; Score 11; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 0.002;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MHAGGDNVSD 161  
 |||||  
 Db 173 MHAGGDNVSD 183

RESULT 2  
 A3893 superoxide dismutase [EC 1.15.1.1] (Cu-Zn) - Brucella abortus  
 C:Species: Brucella abortus  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Aug-2004  
 C:Accession: A3893  
 R:Reck, B.L.; Tabatabai, L.B.; Mayfield, J.E.  
 Biochemistry 29, 372-376, 1990  
 A:Title: A protein isolated from Brucella abortus is a Cu-Zn superoxide dismutase.  
 A:Reference number: A3893; MUID:90148961; PMID:2105741  
 A:Accession: A3893  
 A:Molecule type: protein  
 A:Residues: 1-154 <BEC>  
 A:Cross-references: UNIPROT:PI5453  
 C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C:Superfamily: Superoxide dismutase [Cu-Zn]  
 C:Keywords: copper; metalloprotein; oxidoreductase; zinc  
 F:48-50-73-128/Binding site: copper (His) #status predicted  
 F:55-150/Disulfide bonds: #status predicted  
 F:147/Active site: Arg #status predicted

Query Match 5.0%; Score 9; DB 1; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDRP 163  
 |||||  
 Db 130 GGDNYSDRP 138

RESULT 3  
 A82183 superoxide dismutase [EC 1.15.1.1] (Cu-Zn) [similarity] - Vibrio cholerae (strain N16961)  
 C:Species: Vibrio cholerae  
 C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 16-Aug-2004  
 C:Accession: A82183  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: A82183  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-171 <HEI>  
 A:Cross-references: UNIPROT:Q9KRO3; GB:AE004235; GB:AE003852; NID:g9656082; PIDN:AAF9473  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VC1583  
 A:Map position: 1  
 C:Superfamily: Superoxide dismutase [Cu-Zn]  
 C:Keywords: metalloprotein; oxidoreductase

Query Match 5.0%; Score 9; DB 2; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGGARI 173  
 |||||  
 Db 156 PLGGGGGARI 164

RESULT 4  
 AD3582 superoxide dismutase [EC 1.15.1.1] [imported] - Brucella melitensis (strain 16M)  
 C:Species: Brucella melitensis  
 C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 16-Aug-2004  
 C:Accession: AD3582  
 R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
 A:Reference number: AD3252; PMID:11756688  
 A:Accession: AD3582  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-174 <KUR>  
 A:Cross-references: UNIPROT:P58645; GB:AE008918; PIDN:AAL53823.1; PID:gl17984757; GSPDB:G  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEII0581  
 A:Map position: II  
 C:Superfamily: Superoxide dismutase [Cu-Zn]  
 C:Keywords: oxidoreductase

Query Match 5.0%; Score 9; DB 2; Length 174;

Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDKP 163  
|||||  
Db 150 GGDNYSDKP 158

RESULT 5  
I39485  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Actinobacillus actinomycetemcomitans (fragm  
C;Species: Actinobacillus actinomycetemcomitans  
C;Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I39485  
R;Kroll, J.S.; Langford, P.R.; Wilks, K.E.; Keil, A.D.  
Microbiology 141, 2271-2279, 1995  
A;Title: Bacterial [Cu-Zn]-superoxide dismutase: Phylogenetically distinct from the euka  
A;Reference number: I39485; MUID:96118708; PMID:7496539  
A;Accession: I39485  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-87 <RES>  
A;Cross-references: UNIPROT:Q59081; EMBL:X83122; NID:gl019745; PIDN:CAA58203.1; PID:gl01  
C;Genetics:  
A;Gene: sodC  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Superfamily: superoxide dismutase [Cu-Zn]  
C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
F;18,27,36,39/Binding site: zinc (His, His, Asp) #status predicted

Query Match 4.4%; Score 8; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDN 158  
|||||  
Db 72 MIHAGGDN 79

RESULT 6  
DSFOCL  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - Photobacterium leiognathi  
C;Species: Photobacterium leiognathi  
C;Date: 03-Aug-1984 #sequence\_revision 12-Apr-1996 #text\_change 16-Aug-2004  
C;Accession: A26689; A00519  
R;Steinman, H.M.  
J. Biol. Chem. 262, 1882-1887, 1987  
A;Title: Bacteriocuprein superoxide dismutase of Photobacterium leiognathi. Isolation an  
A;Reference number: A26689; MUID:87109348; PMID:3805055  
A;Accession: A26689  
A;Molecule type: DNA  
A;Residues: 1-173 <STE>  
A;Cross-references: UNIPROT:P00446; GB:J02658; NID:9150710; PIDN:AAA25632.1; PID:gl50711  
R;Steffens, G.J.; Bannister, J.V.; Bannister, W.H.; Flohe, L.; Gunzler, W.A.; Kim, S.M.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 675-690, 1983  
A;Title: The primary structure of Cu-Zn superoxide dismutase from Photobacterium leiogna  
A;Reference number: A00519; MUID:83289129; PMID:6884993  
A;Accession: A00519  
A;Molecule type: protein  
A;Residues: 23-173 <ST2>  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Superfamily: Superoxide dismutase [Cu-Zn]  
C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-173/Product: superoxide dismutase [Cu-Zn] #status predicted <MAT>  
F;67,69,92,147/Binding site: copper (His) #status predicted  
F;74-169/Disulfide bonds: #status predicted  
F;92,101,110,113/Binding site: zinc (His, His, Asp) #status predicted  
F;166/Active site: Arg #status predicted

Query Match 4.4%; Score 8; DB 1; Length 173;

Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDN 158  
|||||  
Db 145 MIHAGGDN 152

RESULT 7  
G85771  
superoxide dismutase precursor (Cu-Zn) [imported] - Escherichia coli (strain O157:H7, su  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: G85771  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: G85771  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-173 <STO>  
A;Cross-references: UNIPROT:P53635; GB:AE005174; NID:gl2515638; PIDN:AAG56635.1; GSPDB:GN  
A;Experimental source: strain O157:H7, substrain BDL933  
C;Genetics:  
C;Superfamily: superoxide dismutase [Cu-Zn]

Query Match 4.4%; Score 8; DB 2; Length 173;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GEHGPHIH 76  
|||||  
Db 62 GEHGPHIH 69

RESULT 8  
JC6004  
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) sodC precursor - Escherichia coli (strain K-12  
C;Species: Escherichia coli  
C;Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 09-Jul-2004  
C;Accession: JC6004; H64921  
R;Imlay, K.R.C.; Imlay, J.A.  
J. Bacteriol. 178, 2564-2571, 1996  
A;Title: Cloning and analysis of sodC, encoding the copper-zinc superoxide dismutase of  
A;Reference number: JC6004; MUID:96196162; PMID:8626323  
A;Accession: JC6004  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-173 <TML>  
A;Cross-references: UNIPROT:P53635; GB:U51242; NID:gl256445; PIDN:AAB03729.1; PID:gl25644  
R;Experimental source: strain K-12  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: H64921  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-173 <BLAT>  
A;Cross-references: GB:AE000259; GB:U00096; NID:gl787921; PIDN:AAC74718.1; PID:gl787934;  
A;Experimental source: strain K-12, substrain MG1655  
C;Comment: This enzyme is a virulence factor secreted into the periplastic space of gram-  
C;Genetics:  
A;Gene: sodC  
A;Map position: 37 min  
C;Function:  
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
C;Superfamily: superoxide dismutase [Cu-Zn]  
C;Keywords: copper; metalloprotein; oxidoreductase; zinc

F1-19/Domain: signal sequence #status predicted <SIG>  
F:20-173/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
F:67-69-92,147/Binding site: copper (His) #status predicted  
F:74-169/Disulfide bonds: #status predicted  
F:92,101,109,112/Binding site: zinc (His, His, Asp) #status predicted  
F:166/Active site: Arg #status predicted

Query Match 4.4%; Score 8; DB 2; Length 173;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GEHGFHIIH 76  
|||||  
Db 62 GEHGFHIIH 69

RESULT 9  
AF0694  
copper-zinc superoxide dismutase [imported] - Salmonella enterica subsp. enterica serovar typhimurium  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 16-Aug-2004  
C:Accession: AF0694  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; et al.  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar typhimurium  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AF0694  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-173 <PAR>  
A:Cross-references: GB:AU513382; PIDN:CAD01927.1; PID:gl6502769; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY1682  
C:Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 4.4%; Score 8; DB 2; Length 173;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GEHGFHIIH 76  
|||||  
Db 62 GEHGFHIIH 69

RESULT 10  
C90923  
superoxide dismutase precursor (Cu-Zn) [imported] - Escherichia coli (strain O157:H7, serotype O157)  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 16-Aug-2004  
C:Accession: C90923  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. et al.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic islands  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: C90923  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-173 <HAY>  
A:Cross-references: UNIPROT:P53635; GB:BA000007; PIDN:BA035778.1; PID:gl13361822; GSPDB:GN00176  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS2355  
C:Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 4.4%; Score 8; DB 2; Length 173;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GEHGFHIIH 76  
|||||  
Db 62 GEHGFHIIH 69

RESULT 11  
H90768  
probable copper/zinc-superoxide dismutase [imported] - Escherichia coli (strain O157:H7, serotype O157)  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: H90768  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. et al.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic islands  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: H90768  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-175 <HAY>  
A:Cross-references: UNIPROT:Q8X9P0; GB:BA000007; PIDN:BA034543.1; PID:gl13360580; GSPDB:GN00176  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS1120  
C:Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 4.4%; Score 8; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172  
|||||  
Db 161 PLGGGGAR 168

RESULT 12  
E90877  
probable copper/zinc-superoxide dismutase [imported] - Escherichia coli (strain O157:H7, serotype O157)  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 16-Aug-2004  
C:Accession: E90877  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. et al.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic islands  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: E90877  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-175 <HAY>  
A:Cross-references: UNIPROT:Q8XGB6; GB:BA000007; PIDN:BA035412.1; PID:gl13361454; GSPDB:GN00176  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS1989  
C:Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 4.4%; Score 8; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172  
|||||  
Db 161 PLGGGGAR 168

RESULT 13  
A41654  
superoxide dismutase (BC 1.15.1.1) (Cu-Zn) precursor - Haemophilus influenzae  
C:Species: Haemophilus influenzae  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Aug-2004  
C:Accession: A41654  
R:Kroll, J.S.; Langford, P.R.; Loynds, B.M. et al.  
J. Bacteriol. 173, 7449-7457, 1991

A;Title: Copper-zinc superoxide dismutase of Haemophilus influenzae and Haemophilus parainfluenzae  
 A;Reference number: A41654; MUID:92041655; PMID:1938942  
 A;Accession: A41654  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-187 <KRO>  
 A;Cross-references: UNIPROT:P25841; GB:M84012; NID:g148881; PIDN:AAA24953.1; PID:g148882  
 C;Genetics:  
 A;Gene: sodC  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: Superoxide dismutase [Cu-Zn]  
 C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
 F;1-35/Domain: signal sequence #status predicted <SIG>  
 F;36-187/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
 F;87-183/Disulfide bonds: #status predicted  
 F;105,114,123,126/Binding site: zinc (His, His, Asp) #status predicted  
 F;180/Active site: Arg #status predicted

Query Match 4.4%; Score 8; DB 1; Length 187;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDN 158  
 |||||  
 Db 159 MIHAGGDN 166

RESULT 14  
 B41654  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - Haemophilus parainfluenzae  
 C;Species: Haemophilus parainfluenzae  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Aug-2004  
 C;Accession: B41654  
 R;Xroll, J.S.; Langford, P.R.; Loynds, B.M.  
 J. Bacteriol. 173, 7449-7457, 1991  
 A;Title: Copper-zinc superoxide dismutase of Haemophilus influenzae and Haemophilus parainfluenzae  
 A;Reference number: A41654; MUID:92041655; PMID:1938942  
 A;Accession: B41654  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-187 <KRO>  
 A;Cross-references: UNIPROT:P25842; GB:M84013; NID:g148883; PIDN:AAA24954.1; PID:g148884  
 C;Genetics:  
 A;Gene: sodC  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: Superoxide dismutase [Cu-Zn]  
 C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
 F;1-35/Domain: signal sequence #status predicted <SIG>  
 F;36-187/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
 F;80,82,105,161/Binding site: copper (His) #status predicted  
 F;87-183/Disulfide bonds: #status predicted  
 F;180/Active site: Arg #status predicted

Query Match 4.4%; Score 8; DB 1; Length 187;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDN 158  
 |||||  
 Db 159 MIHAGGDN 166

RESULT 15  
 JC5718  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) precursor - Haemophilus ducreyi  
 C;Species: Haemophilus ducreyi  
 C;Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 09-Jul-2004  
 C;Accession: JC5718  
 R;Stevens, M.K.; Hassett, D.J.; Radolf, J.D.; Hansen, E.J.  
 Gene 183, 35-40, 1996  
 A;Title: Cloning and sequencing of the gene encoding the Cu,Zn-superoxide dismutase of H. ducreyi  
 A;Reference number: JC5718; MUID:97149276; PMID:8996084  
 A;Accession: JC5718

A;Molecule type: DNA  
 A;Residues: 1-199 <STE>  
 A;Cross-references: UNIPROT:Q59452; GB:U47664; NID:g1305411; PIDN:AAB41293.1; PID:g130541  
 C;Genetics:  
 A;Gene: sodC  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: superoxide dismutase [Cu-Zn]  
 C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
 F;1-22/Domain: signal sequence #status predicted <SIG>  
 F;23-199/Product: superoxide dismutase (Cu-Zn) #status predicted <MAT>  
 F;92,94,117,173/Binding site: copper (His) #status predicted  
 F;99-195/Disulfide bonds: #status predicted  
 F;192/Active site: Arg #status predicted

Query Match 4.4%; Score 8; DB 2; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDN 158  
 |||||  
 Db 171 MIHAGGDN 178

RESULT 16  
 F85741  
 hypothetical protein Z2347 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
 C;Species: Escherichia coli  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Aug-2004  
 C;Accession: F85741  
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, N.; et al.  
 Nature 409, 529-533, 2001  
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A;Reference number: A85480; MUID:21074935; PMID:11206551  
 A;Accession: F85741  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-274 <STO>  
 A;Cross-references: UNIPROT:O8X9P0; GB:AE005174; NID:g12515337; PIDN:AAG56394.1; GSPDB:G12515337  
 A;Experimental source: strain O157:H7, substrain EDL933  
 C;Genetics:  
 A;Gene: Z2347  
 C;Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 4.4%; Score 8; DB 2; Length 274;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172  
 |||||  
 Db 260 PLGGGGAR 267

RESULT 17  
 E85842  
 probable superoxide dismutase Z3312 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
 C;Species: Escherichia coli  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Aug-2004  
 C;Accession: E85842  
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, N.; et al.  
 Nature 409, 529-533, 2001  
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A;Reference number: A85480; MUID:21074935; PMID:11206551  
 A;Accession: E85842  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-328 <STO>  
 A;Cross-references: UNIPROT:O8X6B6; GB:AE005174; NID:g12515337; PIDN:AAG57201.1; GSPDB:G12515337  
 A;Experimental source: strain O157:H7, substrain EDL933  
 C;Genetics:  
 A;Gene: Z3312

C;Superfamily: Superoxide dismutase [Cu-Zn]

Query Match 4.4%; Score 8; DB 2; Length 328;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172

Db 314 PLGGGGAR 321

RESULT 18

C72299

sugar kinase - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C;Accession: C72299

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: C72299

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-476 <ARN>

A;Cross-references: UNIPROT:Q9XOG2; GB:AE001767; GB:AE000512; NID:g4981611; PIDN:AAD3615

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TM1073

C;Superfamily: rhamnulokinase

Query Match

Best Local Similarity 4.4%; Score 8; DB 2; Length 476;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 RLTVKEIK 146

Db 25 RLTVKEIK 32

RESULT 19

JT0531

muscarinic acetylcholine receptor M5 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C;Accession: JT0531; A33354; C37121

R;Bonner, T.I.; Young, A.C.; Brann, M.R.; Buckley, N.J.

Neuron 1, 403-410, 1988

A;Title: Cloning and expression of the human and rat m5 muscarinic acetylcholine recepto

A;Reference number: JT0530; MUID:90166521; PMID:3272174

A;Accession: JT0531

A;Molecule type: DNA

A;Residues: 1-531 <BON>

A;Cross-references: UNIPROT:P08911

R;Liao, C.F.; Themmen, A.P.N.; Joho, R.; Barberis, C.; Birnbaumer, M.; Birnbaumer, L.

J. Biol. Chem. 264, 7328-7337, 1989

A;Title: Molecular cloning and expression of a fifth muscarinic acetylcholine receptor.

A;Reference number: A33354; MUID:89214170; PMID:2540186

A;Accession: A33354

A;Molecule type: DNA

A;Residues: 1-531 <LIA>

A;Cross-references: GB:M22925; NID:g205311; PIDN:AAA41572.1; PID:g205312; GB:J04706

A;Note: the nucleotide sequence for residues 101-120 and the translation 121-140 are not

A;Note: the authors translated the codon CAG for residue 19 as Glu, AAC for residue 65 a

R;Kuttenbach, E.; Curtis, C.A.M.; Pedder, E.K.; Aitken, A.; Harris, A.C.M.; Hulme, E.C.

J. Biol. Chem. 265, 13702-13708, 1990

A;Title: Muscarinic acetylcholine receptors. Peptide sequencing identifies residues invo

A;Reference number: A37121; MUID:90337982; PMID:2380182

A;Accession: C37121

A;Status: preliminary

A;Molecule type: protein

A;Residues: 66-128 <KUR>

C;Comment: Muscarinic acetylcholine receptors mediate many of the actions of the neurotr

C;Superfamily: G protein-coupled rhodopsin

C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phospho

F;29-52/Domain: transmembrane #status predicted <TM1>

F;66-86/Domain: transmembrane #status predicted <TM2>

F;104-125/Domain: transmembrane #status predicted <TM3>

F;146-168/Domain: transmembrane #status predicted <TM4>

F;191-213/Domain: transmembrane #status predicted <TM5>

F;443-463/Domain: transmembrane #status predicted <TM6>

F;478-496/Domain: transmembrane #status predicted <TM7>

F;7,12/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 4.4%; Score 8; DB 2; Length 531;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLNVKADG 129

Db 381 RLNVKADG 388

RESULT 20

JT0530

muscarinic acetylcholine receptor M5 - human

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 21-Jan-2000

C;Accession: JT0530

R;Bonner, T.I.; Young, A.C.; Brann, M.R.; Buckley, N.J.

Neuron 1, 403-410, 1988

A;Title: Cloning and expression of the human and rat m5 muscarinic acetylcholine recepto

A;Reference number: JT0530; MUID:90166521; PMID:3272174

A;Accession: JT0530

A;Molecule type: DNA

A;Residues: 1-532 <BON>

C;Comment: Muscarinic acetylcholine receptors mediate many of the actions of the neurotr

C;Superfamily: G protein-coupled rhodopsin

C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phospho

F;30-53/Domain: transmembrane #status predicted <TM1>

F;67-87/Domain: transmembrane #status predicted <TM2>

F;105-126/Domain: transmembrane #status predicted <TM3>

F;147-169/Domain: transmembrane #status predicted <TM4>

F;192-214/Domain: transmembrane #status predicted <TM5>

F;444-464/Domain: transmembrane #status predicted <TM6>

F;479-498/Domain: transmembrane #status predicted <TM7>

F;8,13/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 4.4%; Score 8; DB 2; Length 532;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 RLNVKADG 129

Db 382 RLNVKADG 389

RESULT 21

AC1716

glyoxalase I homolog lin2271 [imported] - Listeria innocua (strain Clip11262)

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AC1716

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U

Science 294, 849-852, 2001

A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madero, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1716

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-129 <GLA>  
 A;Cross-references: UNIPROT:Q929K4; GB:AL592022; PIDN:CAC97499.1; PID:g16414783; GSPDB:G  
 A;Experimental source: strain Clp11262  
 C;Genetics:  
 A;Gene: lin271

Query Match 3.9%; Score 7; DB 2; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 TDLKGLP 67  
 |||||  
 DB 98 TDLKGLP 104

RESULT 22  
 AHJ345  
 glyoxalase I homolog lmo2168 [imported] - Listeria monocytogenes (strain EGD-e)  
 C;Species: Listeria monocytogenes  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C;Accession: AHJ345  
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
 A;Title: Comparative genomics of Listeria species  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AHJ345  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-129 <GLA>  
 A;Cross-references: UNIPROT:Q9Y5A1; GB:NC\_003210; PIDN:CAD00246.1; PID:g16411638; GSPDB:  
 A;Experimental source: strain EGD-e  
 C;Genetics:  
 A;Gene: lmo2168

Query Match 3.9%; Score 7; DB 2; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 TDLKGLP 67  
 |||||  
 DB 98 TDLKGLP 104

RESULT 23  
 F95926  
 hypothetical exported protein [imported] - Sinorhizobium meliloti (strain 1021) magaplas  
 C;Species: Sinorhizobium meliloti  
 C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
 C;Accession: F95926  
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
 A;Reference number: A95842; MUID:21396508; PMID:11481431  
 A;Accession: F95926  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-140 <KUR>  
 A;Cross-references: UNIPROT:Q92VM2; GB:AL591985; PIDN:CAC49078.1; PID:g15140563; GSPDB:G  
 A;Experimental source: strain 1021, megaplasmid pSymB  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A;Reference number: A96039; MUID:21368234; PMID:11474104  
 A;Contents: annotation  
 C;Genetics:

A;Gene: SMB21099  
 A;Genome: plasmid

Query Match 3.9%; Score 7; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 LTAGLQA 96  
 |||||  
 DB 19 LTAGLQA 25

RESULT 24  
 F70321  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Aquifex aeolicus  
 C;Species: Aquifex aeolicus  
 C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 16-Aug-2004  
 C;Accession: F70321  
 R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove  
 V.  
 Nature 392, 353-358, 1998  
 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A;Reference number: A70300; MUID:98196666; PMID:9537320  
 A;Accession: F70321  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-171 <AQF>  
 A;Cross-references: UNIPROT:O66602; GB:AE000679; NID:g2982936; PIDN:AAC06553.1; PID:g2982  
 A;Experimental source: strain VFS  
 C;Genetics:  
 A;Gene: sodC2  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: Superoxide dismutase [Cu-Zn]  
 C;Keywords: copper; metalloprotein; oxidoreductase; zinc  
 F;164/Active site: Arg #status predicted

Query Match 3.9%; Score 7; DB 2; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 RIACGVI 178  
 |||||  
 DB 164 RIACGVI 170

RESULT 25  
 JE0097  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) 4-1 - Caenorhabditis elegans  
 N;Alternate names: SOD4-1  
 C;Species: Caenorhabditis elegans  
 C;Date: 19-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
 C;Accession: JE0097  
 R;Fuji, M.; Ishii, N.; Joguchi, A.; Yasuda, K.; Ayusawa, D.  
 DNA Res. 5, 25-30, 1998  
 A;Title: A novel superoxide dismutase gene encoding membrane-bound and extracellular iso  
 A;Reference number: JE0097; MUID:98290544; PMID:9628580  
 A;Accession: JE0097  
 A;Molecule type: DNA  
 A;Residues: 1-176 <FUJ>  
 A;Cross-references: UNIPROT:P34461; DDBJ:AB003924; NID:g3135194; PIDN:BAA28262.1; PID:g31  
 C;Comment: This protein is an extracellular form.  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: superoxide dismutase [Cu-Zn]  
 C;Keywords: copper; glycoprotein; metalloprotein; oxidoreductase; zinc  
 F;56/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;70,72,87,144/Binding site: copper (His) #status predicted  
 F;81-170/Disulfide bonds: #status predicted  
 F;87,95,104,107/Binding site: zinc (His, His, Asp) #status predicted  
 F;167/Active site: Arg #status predicted

Query Match 3.9%; Score 7; DB 2; Length 176;



```

Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 HGFHIHE 77
DB 67 HGFHIHE 73

RESULT 26
S40984
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - Caenorhabditis elegans
N:Alternate names: hypothetical protein F55H2.1
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S40984
R:Craxton, M.; Hawkins, T.; Thomas, K.
submitted to the EMBL Data Library, October 1993
A:Reference number: S40984
A:Accession: S40984
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-184 <CRA>
A:Cross-references: UNIPROT:P34461; EMBL:Z27080; NID:g414620; PID:g414621
C:Genetics:
A:Introns: 21/1; 54/3; 114/3; 146/3
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase [Cu-Zn]
C:Keywords: copper; metalloprotein; oxidoreductase; zinc
F:70,72,87,144/Binding site: copper (His) #status predicted
F:81-170/Disulfide bonds: #status predicted
F:87,95,104,107/Binding site: zinc (His, His, Asp) #status predicted
F:167/Active site: Arg #status predicted

Query Match 3.9%; Score 7; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 HGFHIHE 77
DB 67 HGFHIHE 73

RESULT 27
E81855
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) NMA1617 [similarity] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 16-Aug-2004
C:Accession: E81855
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: E81855
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <PAR>
A:Cross-references: UNIPROT:P57005; GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB8484
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: sodC; NMA1617
C:Superfamily: Superoxide dismutase [Cu-Zn]
C:Keywords: metalloprotein; oxidoreductase
F:179/Active site: Arg #status predicted

Query Match 3.9%; Score 7; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 HGFHIHE 77
DB 76 HGFHIHE 82

```

```

RESULT 28
F81088
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) NMB1398 [similarity] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 16-Aug-2004
C:Accession: F81088
R:Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ven
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: F81088
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <TET>
A:Cross-references: UNIPROT:Q59623; GB:AE002488; GB:AE002098; NID:g7226631; PIDN:AAF4176
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1398
C:Superfamily: Superoxide dismutase [Cu-Zn]
C:Keywords: metalloprotein; oxidoreductase
F:179/Active site: Arg #status predicted

Query Match 3.9%; Score 7; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 HGFHIHE 77
DB 76 HGFHIHE 82

RESULT 29
JE0098
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) 4-2 - Caenorhabditis elegans
N:Alternate names: SOD4-2
C:Species: Caenorhabditis elegans
C:Date: 19-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: JE0098
R:Fujii, M.; Ishii, N.; Joguchi, A.; Yasuda, K.; Ayusawa, D.
DNA Res. 5, 25-30, 1998
A:Title: A novel superoxide dismutase gene encoding membrane-bound and extracellular iso
A:Reference number: JE0097; MUID:98290544; PMID:9628580
A:Accession: JE0098
A:Molecule type: DNA
A:Residues: 1-221 <FOU>
A:Cross-references: UNIPROT:P34461; UNIPROT:Q27538; DDBJ:AB003924
C:Comment: This protein is a membrane-bound form.
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase [Cu-Zn]
C:Keywords: copper; glycoprotein; metalloprotein; oxidoreductase; zinc
F:201-221/Domain: transmembrane #status predicted <TMM>
F:56/Binding site: carboxylate (Asn) (covalent) #status predicted
F:70,72,87,144/Binding site: copper (His) #status predicted
F:81-170/Disulfide bonds: #status predicted
F:87,95,104,107/Binding site: zinc (His, His, Asp) #status predicted
F:167/Active site: Arg #status predicted

Query Match 3.9%; Score 7; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 HGFHIHE 77
DB 67 HGFHIHE 73

RESULT 30

```



## F90626

ATP synthase F0 chain 6 [imported] - Eudromia elegans mitochondrion  
C;Species: mitochondrion Eudromia elegans  
C;Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 09-Jul-2004  
C;Accession: F90626  
R;Haddrath, O.; Baker, A.J.  
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001  
A;Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene  
A;Reference number: A99613; MUID:21263106; PMID:11370967  
A;Accession: F90626  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-227 <KUR>  
A;Cross-references: UNIPROT:Q9B6S9; GB:NC\_002772; NID:gl14141819; PIDN:NP\_115278.1; GSPDB  
C;Genetics:  
A;Gene: ATP6  
A;Genome: mitochondrion  
A;Genetic code: SGC1  
C;Superfamily: H+-transporting ATP synthase protein 6  
C;Keywords: mitochondrion

Query Match 3.9%; Score 7; DB 2; Length 227;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TISLLTS 19

Db 188 TISLLTS 194

## RESULT 31

AE2049  
ATP-binding protein of ABC transporter all1947 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AE2049  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AE2049  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-232 <KUR>  
A;Cross-references: UNIPROT:Q9YVM7; GB:BA000019; PIDN:BA073646.1; PID:gl7131037; GSPDB:G  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all1947  
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 3.9%; Score 7; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 PRLTVKE 144

Db 89 PRLTVKE 95

## RESULT 32

T30864  
stbB protein - Salmonella typhimurium plasmid pKM101  
C;Species: Salmonella typhimurium  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T30864  
R;Winans, S.C.  
submitted to the EMBL Data Library, January 1998  
A;Reference number: Z20310  
A;Accession: T30864  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA

A;Residues: 1-238 <WIN>  
A;Cross-references: UNIPROT:P97109; EMBL:U43676; NID:g2801367; PID:g2801371; PIDN:AAB9726  
C;Genetics:  
A;Genome: plasmid pKM101  
C;Superfamily: Salmonella typhimurium plasmid pKM101 stbB protein

Query Match 3.9%; Score 7; DB 2; Length 238;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 LLAPRLT 141

Db 21 LLAPRLT 27

## RESULT 33

S45407  
probable membrane protein YBL095w - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein YBL0835  
C;Species: Saccharomyces cerevisiae  
C;Date: 09-Jun-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Jul-2004  
C;Accession: S45407; S45836; S59204  
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.  
submitted to the EMBL Data Library, May 1994  
A;Description: Sequence analysis of a 78,6 kb segment of the left end of Saccharomyces cer  
A;Reference number: S45387  
A;Accession: S45407  
A;Molecule type: DNA  
A;Residues: 1-270 <OBE>  
A;Cross-references: UNIPROT:P38172; EMBL:X79489; NID:g496661; PID:g496680  
A;Experimental source: strain S288C  
R;Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.  
submitted to the Protein Sequence Database, August 1994  
A;Reference number: S45816  
A;Accession: S45836  
A;Molecule type: DNA  
A;Residues: 1-270 <DOM>  
A;Cross-references: EMBL:Z35857; NID:g536155; PID:g536157; MIPS:YBL095w  
A;Experimental source: strain S288C  
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.  
Yeast 11, 1103-1112, 1995  
A;Title: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cerevisi  
A;Reference number: S59184; MUID:96076635; PMID:7502586  
A;Accession: S59204  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-270 <OBW>  
A;Cross-references: EMBL:X79489; NID:g496661; PIDN:CAA56007.1; PID:g496680  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994  
C;Genetics:  
A;Cross-references: SGD:S0000191  
A;Map position: 2L  
C;Keywords: transmembrane protein  
F;15-31/Domain: transmembrane #status predicted <TM1>  
F;153-169/Domain: transmembrane #status predicted <TM2>

Query Match 3.9%; Score 7; DB 2; Length 270;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 VKEIKGR 148

Db 200 VKEIKGR 206

## RESULT 34

F86171  
hypothetical protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cross)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: F86171  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: F86171  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-271 <STO>  
 A:Cross-references: UNIPROT:Q9ZWC4; GB:AE005172; NID:G4204285; PIDN:AAD10666.1; GSPDB:GN  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: vegetative storage protein; glucose-6-phosphatase catalytic domain homolog

Query Match 3.9%; Score 7; DB 2; Length 271;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 IKTDLKG 65  
 Db 47 IKTDLKG 53  
 |||||

RESULT 35  
 A:Accession: A82661  
 conserved hypothetical protein Atu0691 [imported] - Agrobacterium tumefaciens (strain C58)  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 A:Accession: A82661  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: A82577; MUID:21608550; PMID:11743193  
 A:Accession: A82661  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-308 <KUR>  
 A:Cross-references: UNIPROT:Q8UHJ1; GB:AE008688; PIDN:AAL41707.1; PID:G17739054; GSPDB:G  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu0691  
 A:Map position: circular chromosome

Query Match 3.9%; Score 7; DB 2; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGAR 172  
 Db 35 LGGGGAR 41  
 |||||

RESULT 36  
 C97443  
 hypothetical protein AGR\_C\_1245 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C:Accession: C97443  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: C97443  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-311 <KUR>  
 A:Cross-references: UNIPROT:Q8UHJ1; GB:AE007869; PIDN:AAK86500.1; PID:G15155652; GSPDB:G  
 C:Genetics:  
 A:Gene: AGR\_C\_1245  
 A:Map position: circular chromosome

Query Match 3.9%; Score 7; DB 2; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGAR 172  
 Db 38 LGGGGAR 44  
 |||||

RESULT 37  
 B75374  
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 A:Accession: B75374  
 R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I  
 S.; Smith, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mal  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: B75374  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-341 <WHI>  
 A:Cross-references: UNIPROT:Q9RTY3; GB:AE002005; GB:AE000513; NID:G6459377; PIDN:AAF1117  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR1616  
 A:Map position: 1

Query Match 3.9%; Score 7; DB 2; Length 341;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 LGGGGAR 172  
 Db 137 LGGGGAR 143  
 |||||

RESULT 38  
 E90057  
 hypothetical protein SA2320 [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 A:Accession: E90057  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: E90057  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-347 <KUR>  
 A:Cross-references: UNIPROT:Q99RA3; GB:BA000018; PID:G13702482; PIDN:BAB43623.1; GSPDB:G  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: SA2320

Query Match 3.9%; Score 7; DB 2; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VTSIVTI 14

db 300 VTSIVTI 306  
|||||

## RESULT 39

S78285

conserved hypothetical protein 382 - Odontella sinensis chloroplast

C;Species: chloroplast Odontella sinensis

C;Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 09-Jul-2004

C;Accession: S78285

R;Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A;Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sinensis

A;Reference number: S78238

A;Accession: S78285

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-382 &lt;KOW&gt;

A;Cross-references: UNIPROT:P49539; EMBL:Z67753; NID:g1185127; PIDN:CAA91658.1; PID:g118

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

C;Genetics:

A;Gene: ycf44

A;Genome: chloroplast

C;Superfamily: conserved hypothetical protein slr2087

C;Keywords: chloroplast

Query Match 3.9%; Score 7; DB 2; Length 382;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 GLQIKTD 62

|||||

DB 357 GLQIKTD 363

## RESULT 40

C87478

hypothetical protein CC1848 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C;Accession: C87478

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: C87478

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-383 &lt;STO&gt;

A;Cross-references: UNIPROT:Q9A778; GB:AE005673; NID:g13423289; PIDN:AAK23823.1; GSPDB:G

C;Genetics:

A;Gene: CC1848

Query Match 3.9%; Score 7; DB 2; Length 383;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 LLAPRLT 141

|||||

DB 174 LLAPRLT 180

Search completed: October 26, 2004, 10:03:50

Job time : 63 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 09:46:30 ; Search time 195 Seconds  
(without alignments)  
531.115 Million cell updates/sec

Title: US-10-009-916A-1

Perfect score: 180

Sequence: 1 MKIKLFFVTSIVTISLLTSI.....DKLPLGGGARIACGVIPN 180

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Uniprot\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID            | Description         |
|------------|-------|-------------|--------|---------------|---------------------|
| 1          | 14    | 7.8         | 174    | 2 Q9MC02      | Q9mc02 phage fcls   |
| 2          | 14    | 7.8         | 174    | 2 Q8ZQF7      | Q8zqf7 salmonella   |
| 3          | 13    | 7.2         | 175    | 2 Q6G3C6      | Q6g3c6 bartonella   |
| 4          | 12    | 6.7         | 194    | 2 Q8ELY4      | Q8ely4 oceanobacil  |
| 5          | 11    | 6.1         | 170    | 2 Q83A08      | Q83a08 coxiella bu  |
| 6          | 11    | 6.1         | 185    | 1 SODC FRATU  | Q59448 francisella  |
| 7          | 11    | 6.1         | 185    | 2 Q8DLA0      | Q8dlao yersinia pe  |
| 8          | 11    | 6.1         | 201    | 2 Q8ZBN3      | Q8zbn3 yersinia pe  |
| 9          | 11    | 6.1         | 201    | 2 AAS60586    | Aas60586 yersinia   |
| 10         | 9     | 5.0         | 154    | 1 SODC BRUAB  | P15453 bruceella ab |
| 11         | 9     | 5.0         | 170    | 2 Q7MFM9      | Q7mfm9 vibrio vuln  |
| 12         | 9     | 5.0         | 170    | 2 Q8DA54      | Q8da54 vibrio vuln  |
| 13         | 9     | 5.0         | 171    | 2 Q9KKQ3      | Q9kkr3 vibrio chol  |
| 14         | 9     | 5.0         | 174    | 1 SODC BRUME  | P58645 bruceella me |
| 15         | 9     | 5.0         | 177    | 1 SODI SALTY  | P53636 salmonella   |
| 16         | 9     | 5.0         | 177    | 2 Q70457      | Q70457 salmonella   |
| 17         | 9     | 5.0         | 177    | 2 CAF06531    | Caf06531 salmonell  |
| 18         | 9     | 5.0         | 189    | 2 Q6MR06      | Q6mr06 bdellovibri  |
| 19         | 9     | 5.0         | 189    | 2 CAE77952    | CAe77952 bdellovib  |
| 20         | 8     | 4.4         | 87     | 1 SODC ACTAC  | Q59081 actinobacil  |
| 21         | 8     | 4.4         | 165    | 2 Q8PDZ3      | Q8pdz3 xanthomonas  |
| 22         | 8     | 4.4         | 165    | 2 Q8PQW1      | Q8pqw1 xanthomonas  |
| 23         | 8     | 4.4         | 171    | 2 Q87G06      | Q87g06 vibrio para  |
| 24         | 8     | 4.4         | 173    | 1 SODC ECOOLI | P53635 escherichia  |
| 25         | 8     | 4.4         | 173    | 1 SODC PHOLE  | P00446 photobacter  |
| 26         | 8     | 4.4         | 173    | 1 SODC SALTY  | Q68901 salmonella   |
| 27         | 8     | 4.4         | 173    | 2 Q751I8      | Q751i8 ashbya goss  |
| 28         | 8     | 4.4         | 173    | 2 Q704S6      | Q704s6 salmonella   |
| 29         | 8     | 4.4         | 173    | 2 Q8Z6P6      | Q8z6p6 salmonella   |
| 30         | 8     | 4.4         | 173    | 2 AAS54170    | Aas54170 ashbya go  |
| 31         | 8     | 4.4         | 173    | 2 CAF06532    | Caf06532 salmonell  |

|     |   |     |     |              |                     |
|-----|---|-----|-----|--------------|---------------------|
| 32  | 8 | 4.4 | 175 | 2 Q7AEAI     | Q7aeal escherichia  |
| 33  | 8 | 4.4 | 175 | 2 Q7AFX5     | Q7afx5 escherichia  |
| 34  | 8 | 4.4 | 178 | 2 Q7UCE0     | Q7uce0 shigella fl  |
| 35  | 8 | 4.4 | 187 | 1 SODC_HAEIN | P25841 haemophilus  |
| 36  | 8 | 4.4 | 187 | 1 SODC_HAEPA | P25842 haemophilus  |
| 37  | 8 | 4.4 | 187 | 2 Q714V4     | Q714v4 haemophilus  |
| 38  | 8 | 4.4 | 187 | 2 AAQ12654   | AAq12654 haemophil  |
| 39  | 8 | 4.4 | 190 | 2 Q83RB4     | Q83rb4 shigella fl  |
| 40  | 8 | 4.4 | 190 | 2 Q8FH80     | Q8fh80 escherichia  |
| 41  | 8 | 4.4 | 199 | 1 SODC_HAEDU | Q59452 haemophilus  |
| 42  | 8 | 4.4 | 233 | 2 Q9CSB6     | Q9csb6 mus musculu  |
| 43  | 8 | 4.4 | 242 | 2 Q9CS22     | Q9cs22 mus musculu  |
| 44  | 8 | 4.4 | 248 | 2 Q8DUA2     | Q8dua2 streptococc  |
| 45  | 8 | 4.4 | 274 | 2 Q8X9P0     | Q8x9p0 escherichia  |
| 46  | 8 | 4.4 | 291 | 2 Q6NCV8     | Q6ncv8 rhodospseudo |
| 47  | 8 | 4.4 | 291 | 2 CAE25805   | CAe25805 rhodopseu  |
| 48  | 8 | 4.4 | 317 | 2 Q941W8     | Q941w8 oryza sativ  |
| 49  | 8 | 4.4 | 328 | 2 Q8X6B6     | Q8x6b6 escherichia  |
| 50  | 8 | 4.4 | 354 | 2 Q9TLR8     | Q9tlr8 cyanidium c  |
| 51  | 8 | 4.4 | 369 | 2 Q6L4Z5     | Q6l4z5 oryza sativ  |
| 52  | 8 | 4.4 | 369 | 2 AAT39179   | Aat39179 oryza sat  |
| 53  | 8 | 4.4 | 369 | 2 AAT39186   | Aat39186 oryza sat  |
| 54  | 8 | 4.4 | 393 | 2 Q9L9L6     | Q9l9l6 pasteurella  |
| 55  | 8 | 4.4 | 419 | 2 Q83FJ8     | Q83fj8 tropheryma   |
| 56  | 8 | 4.4 | 419 | 2 Q83HA2     | Q83ha2 tropheryma   |
| 57  | 8 | 4.4 | 476 | 2 Q9X0G2     | Q9x0g2 thermotoga   |
| 58  | 8 | 4.4 | 528 | 2 Q9PTF6     | Q9ptf6 gallus gall  |
| 59  | 8 | 4.4 | 531 | 1 ACM5_RAT   | P08911 rattus norv  |
| 60  | 8 | 4.4 | 532 | 1 ACM5_HUMAN | P08912 homo sapien  |
| 61  | 8 | 4.4 | 532 | 1 ACM5_MACMU | P56490 macaca mula  |
| 62  | 8 | 4.4 | 532 | 2 Q6NUM3     | Q6num3 homo sapien  |
| 63  | 8 | 4.4 | 532 | 2 Q8IWM0     | Q8iwm0 homo sapien  |
| 64  | 8 | 4.4 | 532 | 2 Q920H4     | Q920h4 mus musculu  |
| 65  | 8 | 4.4 | 532 | 2 AAH68528   | AAh68528 homo sapi  |
| 66  | 8 | 4.4 | 539 | 2 Q7S7G8     | Q7s7g8 neurospora   |
| 67  | 8 | 4.4 | 539 | 2 CAE76133   | CAe76133 neurospor  |
| 68  | 8 | 4.4 | 575 | 2 Q8VEJ8     | Q8vej8 mus musculu  |
| 69  | 8 | 4.4 | 733 | 2 Q926C7     | Q926c7 rhizobium m  |
| 70  | 8 | 4.4 | 750 | 2 Q8CGC6     | Q8cgc6 mus musculu  |
| 71  | 8 | 4.4 | 759 | 2 Q9NW13     | Q9nw13 homo sapien  |
| 72  | 8 | 4.4 | 759 | 2 Q96CV3     | Q96cv3 homo sapien  |
| 73  | 8 | 4.4 | 956 | 2 Q9UCV3     | Q9ucv3 homo sapien  |
| 74  | 7 | 3.9 | 81  | 2 P84118     | P84118 periplaneta  |
| 75  | 7 | 3.9 | 83  | 2 Q854G8     | Q854g8 mycobacteri  |
| 76  | 7 | 3.9 | 106 | 2 Q7RZW1     | Q7rzw1 neurospora   |
| 77  | 7 | 3.9 | 113 | 2 Q978H1     | Q978h1 thermoplasma |
| 78  | 7 | 3.9 | 120 | 2 Q9L224     | Q9l224 streptomyce  |
| 79  | 7 | 3.9 | 124 | 2 Q836H2     | Q836h2 enterococcu  |
| 80  | 7 | 3.9 | 129 | 2 Q929K4     | Q929k4 listeria in  |
| 81  | 7 | 3.9 | 129 | 2 Q8Y5A1     | Q8y5a1 listeria mo  |
| 82  | 7 | 3.9 | 129 | 2 Q71XJ8     | Q71xj8 listeria mo  |
| 83  | 7 | 3.9 | 129 | 2 AAT04967   | Aat04967 listeria   |
| 84  | 7 | 3.9 | 132 | 2 Q6YSY2     | Q6ysy2 oryza sativ  |
| 85  | 7 | 3.9 | 132 | 2 BAC84758   | Bac84758 oryza sat  |
| 86  | 7 | 3.9 | 133 | 2 Q72JN1     | Q72jn1 thermus the  |
| 87  | 7 | 3.9 | 133 | 2 AAS81087   | Aas81087 thermus t  |
| 88  | 7 | 3.9 | 134 | 2 Q9ATV9     | Q9atv9 shigella fl  |
| 89  | 7 | 3.9 | 134 | 2 AAT9022    | Aat9022 shigella    |
| 90  | 7 | 3.9 | 138 | 2 Q9AFR2     | Q9afr2 shigella fl  |
| 91  | 7 | 3.9 | 138 | 2 Q7BCL0     | Q7bcl0 shigella fl  |
| 92  | 7 | 3.9 | 140 | 2 Q92VM2     | Q92vm2 rhizobium m  |
| 93  | 7 | 3.9 | 146 | 2 Q75NJ1     | Q75nj1 saccharopol  |
| 94  | 7 | 3.9 | 146 | 2 RAD15073   | Rad15073 saccharop  |
| 95  | 7 | 3.9 | 151 | 1 SODC_HAIRO | P81326 halocynthia  |
| 96  | 7 | 3.9 | 153 | 2 Q897Y0     | Q897y0 clostridium  |
| 97  | 7 | 3.9 | 154 | 2 Q6SA03     | Q6sa03 bombyx mori  |
| 98  | 7 | 3.9 | 154 | 2 AAR97568   | Aar97568 bombyx mo  |
| 99  | 7 | 3.9 | 159 | 2 Q851G3     | Q851g3 oryza sativ  |
| 100 | 7 | 3.9 | 160 | 2 Q6ZB13     | Q6zb13 oryza sativ  |
| 101 | 7 | 3.9 | 160 | 2 Q7TLX4     | Q7tlx4 choristoneu  |
| 102 | 7 | 3.9 | 160 | 2 BAD01286   | Bad01286 oryza sat  |
| 103 | 7 | 3.9 | 161 | 2 Q8V9Y1     | Q8v9y1 choristoneu  |
| 104 | 7 | 3.9 | 162 | 1 SODC_LBGP  | P53637 legionella   |

|     |   |     |     |   |             |                     |     |   |     |     |   |            |                     |
|-----|---|-----|-----|---|-------------|---------------------|-----|---|-----|-----|---|------------|---------------------|
| 105 | 7 | 3.9 | 163 | 2 | Q7QJF8      | Q7qf8 anopheles g   | 178 | 7 | 3.9 | 311 | 2 | Q7D0X2     | Q7d0x2 agrobacteri  |
| 106 | 7 | 3.9 | 170 | 2 | Q7VWV9      | Q7vwv9 bordetella   | 179 | 7 | 3.9 | 313 | 2 | Q6BFQ7     | Q6bfq7 paramecium   |
| 107 | 7 | 3.9 | 170 | 2 | Q7WH53      | Q7wh53 bordetella   | 180 | 7 | 3.9 | 317 | 2 | Q8A5N5     | Q8a5n5 bacteroides  |
| 108 | 7 | 3.9 | 171 | 1 | SOD2_AQUAE  | O66602 aquifex aso  | 181 | 7 | 3.9 | 320 | 2 | Q8DMV2     | Q8dmv2 synecococc   |
| 109 | 7 | 3.9 | 172 | 2 | Q6ND84      | Q6nd84 rhodopseudo  | 182 | 7 | 3.9 | 321 | 2 | Q74FI7     | Q74fi7 geobacter s  |
| 110 | 7 | 3.9 | 172 | 2 | Q7P1D2      | Q7p1d2 chromobacte  | 183 | 7 | 3.9 | 321 | 2 | AAR33952   | Aar33952 geobacter  |
| 111 | 7 | 3.9 | 172 | 2 | CAB25669    | Cae25669 rhodopseu  | 184 | 7 | 3.9 | 322 | 2 | Q8JFR3     | Q8jfr3 brachydanio  |
| 112 | 7 | 3.9 | 174 | 2 | Q8TN71      | Q8tn71 methanosarc  | 185 | 7 | 3.9 | 323 | 1 | VATC_THETH | P74902 thermus the  |
| 113 | 7 | 3.9 | 174 | 2 | Q6P531      | Q6p531 erwinia car  | 186 | 7 | 3.9 | 323 | 2 | Q72J70     | Q72j70 thermus t    |
| 114 | 7 | 3.9 | 179 | 2 | Q6KTI6      | Q6kti6 bacillus ce  | 187 | 7 | 3.9 | 323 | 2 | AAS81253   | Aas81253 thermus t  |
| 115 | 7 | 3.9 | 179 | 2 | Q6HBZ8      | Q6hbz8 bacillus th  | 188 | 7 | 3.9 | 327 | 2 | Q7R765     | Q7r765 plasmidium   |
| 116 | 7 | 3.9 | 179 | 2 | Q72YH5      | Q72yh5 bacillus ce  | 189 | 7 | 3.9 | 329 | 2 | Q6MBN2     | Q6mbn2 parachlamy   |
| 117 | 7 | 3.9 | 179 | 2 | Q7NPG6      | Q7npg6 gloeobacter  | 190 | 7 | 3.9 | 329 | 2 | CAR24017   | Car24017 parachlam  |
| 118 | 7 | 3.9 | 179 | 2 | Q816F2      | Q816f2 bacillus ce  | 191 | 7 | 3.9 | 332 | 2 | Q893K5     | Q893k5 clostridium  |
| 119 | 7 | 3.9 | 179 | 2 | Q81K66      | Q81k66 bacillus an  | 192 | 7 | 3.9 | 332 | 2 | Q9RTY3     | Q9rty3 deinococcus  |
| 120 | 7 | 3.9 | 179 | 2 | AAR85496    | Aar85496 bacillus   | 193 | 7 | 3.9 | 341 | 2 | Q8HNS5     | Q8hns5 cercaritetus |
| 121 | 7 | 3.9 | 179 | 2 | AAS43947    | Aas43947 bacillus   | 194 | 7 | 3.9 | 347 | 2 | Q8HJ3      | Q8hj3 cercaritetus  |
| 122 | 7 | 3.9 | 179 | 2 | AAT34268    | Aat34268 bacillus   | 195 | 7 | 3.9 | 347 | 2 | Q6G6E2     | Q6g6e2 staphylococ  |
| 123 | 7 | 3.9 | 181 | 2 | Q9V5Z3      | Q9v5z3 drosophila   | 196 | 7 | 3.9 | 347 | 2 | Q8GDR6     | Q8gdr6 staphylococ  |
| 124 | 7 | 3.9 | 181 | 2 | AAF58647    | Aaf58647 drosophil  | 197 | 7 | 3.9 | 347 | 2 | Q8NUS6     | Q8nus6 staphylococ  |
| 125 | 7 | 3.9 | 183 | 1 | SODC_HAECCO | P51547 haemochus    | 198 | 7 | 3.9 | 347 | 2 | Q99RA3     | Q99ra3 staphylococ  |
| 126 | 7 | 3.9 | 186 | 1 | SODC_NEIMA  | P57005 neisseria m  | 199 | 7 | 3.9 | 347 | 2 | Q7A3H0     | Q7a3h0 staphylococ  |
| 127 | 7 | 3.9 | 186 | 1 | SODC_NEIME  | Q59623 neisseria m  | 200 | 7 | 3.9 | 351 | 2 | Q6ZBS0     | Q6zbs0 oryza sativ  |
| 128 | 7 | 3.9 | 188 | 2 | Q6HF7       | Q6hf7 bacillus th   | 201 | 7 | 3.9 | 351 | 2 | BAD03047   | Bad03047 oryza sat  |
| 129 | 7 | 3.9 | 188 | 2 | Q8KL00      | Q8kl00 rhizobium e  | 202 | 7 | 3.9 | 351 | 2 | BAD03243   | Bad03243 oryza sat  |
| 130 | 7 | 3.9 | 190 | 1 | HIS7_STRMU  | Q8dtq9 streptococ   | 203 | 7 | 3.9 | 354 | 2 | Q97C19     | Q97c19 thermoplasm  |
| 131 | 7 | 3.9 | 194 | 2 | Q7W9K6      | Q7w9k6 bordetella   | 204 | 7 | 3.9 | 369 | 1 | LGT_MYCGA  | Q7nae3 mycoplasma   |
| 132 | 7 | 3.9 | 194 | 2 | Q6WAB6      | Q6wab6 parachlamy   | 205 | 7 | 3.9 | 371 | 2 | Q8GHQ5     | Q8ghq5 pseudomonas  |
| 133 | 7 | 3.9 | 205 | 2 | CAP24483    | Caf24483 parachlam  | 206 | 7 | 3.9 | 372 | 2 | Q6QR95     | Q6qr95 eurycea spe  |
| 134 | 7 | 3.9 | 205 | 2 | Q7Q9H5      | Q7q9h5 anopheles g  | 207 | 7 | 3.9 | 372 | 2 | AAS90818   | Aas90818 eurycea s  |
| 135 | 7 | 3.9 | 206 | 2 | Q6OVQ5      | Q6ovq5 anopheles g  | 208 | 7 | 3.9 | 373 | 2 | Q8QB23     | Q8qeb23 shewanella  |
| 136 | 7 | 3.9 | 211 | 2 | AAS17758    | Aas17758 anopheles  | 209 | 7 | 3.9 | 374 | 2 | Q8EC17     | Q8ec17 shewanella   |
| 137 | 7 | 3.9 | 211 | 2 | Q8L713      | Q8l713 arabidopsis  | 210 | 7 | 3.9 | 376 | 2 | Q8SX29     | Q8sx29 drosophila   |
| 138 | 7 | 3.9 | 212 | 2 | AAT41859    | Aat41859 arabidops  | 211 | 7 | 3.9 | 382 | 1 | YC44_ODOSI | P49539 odontella s  |
| 139 | 7 | 3.9 | 212 | 2 | Q7VXX8      | Q7vxx8 bordetella   | 212 | 7 | 3.9 | 383 | 2 | Q8RFZ2     | Q8rfz2 bruceella su |
| 140 | 7 | 3.9 | 213 | 2 | Q8MUS1      | Q8mus1 anopheles g  | 213 | 7 | 3.9 | 383 | 2 | Q98HW5     | Q98hw5 rhizobium l  |
| 141 | 7 | 3.9 | 214 | 2 | Q7VXX8      | Q7vxx8 bordetella   | 214 | 7 | 3.9 | 383 | 2 | Q9A778     | Q9a778 caulobacter  |
| 142 | 7 | 3.9 | 216 | 2 | Q7WI95      | Q7wi95 bordetella   | 215 | 7 | 3.9 | 383 | 2 | Q8YIF9     | Q8yif9 bruceella me |
| 143 | 7 | 3.9 | 221 | 1 | SODE_CAEEL  | P34461 caenorhabdi  | 216 | 7 | 3.9 | 390 | 2 | Q8GF56     | Q8gff6 dictyosteli  |
| 144 | 7 | 3.9 | 221 | 2 | Q72F06      | Q72f06 desulfovibr  | 217 | 7 | 3.9 | 391 | 2 | Q72S10     | Q72s10 leptospira   |
| 145 | 7 | 3.9 | 221 | 2 | AAS94895    | Aas94895 desulfovi  | 218 | 7 | 3.9 | 391 | 2 | AAS70173   | Aas70173 leptospir  |
| 146 | 7 | 3.9 | 225 | 2 | Q84ZK3      | Q84zr3 oryza sativ  | 219 | 7 | 3.9 | 394 | 2 | Q7PQ57     | Q7pq57 anopheles g  |
| 147 | 7 | 3.9 | 227 | 2 | Q9B6S9      | Q9b6s9 eudromia el  | 220 | 7 | 3.9 | 399 | 2 | Q6Z128     | Q6z128 oryza sativ  |
| 148 | 7 | 3.9 | 230 | 2 | Q89E09      | Q89e09 bradyrhizob  | 221 | 7 | 3.9 | 399 | 2 | BAD15490   | Bad15490 oryza sat  |
| 149 | 7 | 3.9 | 232 | 2 | Q8YVM7      | Q8yvm7 anabaena sp  | 222 | 7 | 3.9 | 402 | 2 | Q8RTH6     | Q8rth6 vibrio chol  |
| 150 | 7 | 3.9 | 233 | 2 | Q7NJ51      | Q7nj51 gloeobacter  | 223 | 7 | 3.9 | 406 | 2 | P93487     | P93487 pisum sativ  |
| 151 | 7 | 3.9 | 233 | 2 | Q7VBC6      | Q7vbc6 prochloroco  | 224 | 7 | 3.9 | 407 | 2 | Q8BL79     | Q8bl79 mus musculu  |
| 152 | 7 | 3.9 | 233 | 2 | Q7W6C6      | Q7w6c6 bordetella   | 225 | 7 | 3.9 | 409 | 2 | Q6SKD6     | Q6skd6 bacillus th  |
| 153 | 7 | 3.9 | 234 | 2 | Q93JL6      | Q93jl6 streptomyce  | 226 | 7 | 3.9 | 409 | 2 | Q73A99     | Q73a99 bacillus ce  |
| 154 | 7 | 3.9 | 238 | 2 | P97109      | P97109 plasmid pkm  | 227 | 7 | 3.9 | 409 | 2 | Q81F53     | Q81f53 bacillus ce  |
| 155 | 7 | 3.9 | 238 | 2 | Q79CA5      | Q79ca5 salmonella   | 228 | 7 | 3.9 | 409 | 2 | Q81S62     | Q81s62 bacillus an  |
| 156 | 7 | 3.9 | 238 | 2 | Q79SD4      | Q79sd4 incn plasm   | 229 | 7 | 3.9 | 409 | 2 | AAS40807   | Aas40807 bacillus   |
| 157 | 7 | 3.9 | 244 | 2 | Q7TGX1      | Q7t6x1 mimivirus.   | 230 | 7 | 3.9 | 409 | 2 | AAT30922   | Aat30922 bacillus   |
| 158 | 7 | 3.9 | 245 | 2 | Q7QCE9      | Q7qce9 anopheles g  | 231 | 7 | 3.9 | 412 | 2 | Q6MY8      | Q6my8 bdellovibri   |
| 159 | 7 | 3.9 | 250 | 2 | Q8DJU8      | Q8dju8 synecococc   | 232 | 7 | 3.9 | 412 | 2 | CAE80775   | Caes80775 bdellovib |
| 160 | 7 | 3.9 | 253 | 1 | HAD_XANAU   | Q60099 xanthobacte  | 233 | 7 | 3.9 | 413 | 2 | Q73P03     | Q73p03 treponema    |
| 161 | 7 | 3.9 | 255 | 2 | Q6LYV1      | Q6lyv1 methanococc  | 234 | 7 | 3.9 | 421 | 2 | AAS11487   | Aas11487 treponema  |
| 162 | 7 | 3.9 | 255 | 2 | CAR30442    | Caf30442 methanoco  | 235 | 7 | 3.9 | 421 | 2 | Q9HHV4     | Q9shv4 halobacteri  |
| 163 | 7 | 3.9 | 270 | 1 | YBJ5_YEAST  | P318172 saccharomyc | 236 | 7 | 3.9 | 421 | 2 | Q9S4G7     | Q9s4g7 leptospira   |
| 164 | 7 | 3.9 | 270 | 2 | AAS56891    | Aas56891 saccharom  | 237 | 7 | 3.9 | 422 | 2 | Q6TNY3     | Q6tny3 galdieria s  |
| 165 | 7 | 3.9 | 271 | 2 | Q9ZWC4      | Q9zwc4 arabidopsis  | 238 | 7 | 3.9 | 422 | 2 | AAR30294   | Aar30294 galdieria  |
| 166 | 7 | 3.9 | 289 | 2 | Q6HJUR9     | Q6hjr9 bacillus th  | 239 | 7 | 3.9 | 424 | 2 | Q7X366     | Q7x366 uncultured   |
| 167 | 7 | 3.9 | 289 | 2 | Q739L1      | Q739l1 bacillus ce  | 240 | 7 | 3.9 | 424 | 2 | Q72ET0     | Q72et0 desulfovibr  |
| 168 | 7 | 3.9 | 289 | 2 | Q81R15      | Q81ri15 bacillus an | 241 | 7 | 3.9 | 424 | 2 | AAS94971   | Aas94971 desulfovi  |
| 169 | 7 | 3.9 | 289 | 2 | AAS41050    | Aas41050 bacillus   | 242 | 7 | 3.9 | 426 | 2 | Q86YU8     | Q86yu8 homo sapien  |
| 170 | 7 | 3.9 | 289 | 2 | AAT31176    | Aat31176 bacillus   | 243 | 7 | 3.9 | 426 | 2 | Q7NGV8     | Q7ngv8 gloeobacter  |
| 171 | 7 | 3.9 | 294 | 2 | Q7QDP9      | Q7qdp9 anopheles g  | 244 | 7 | 3.9 | 429 | 2 | Q7X2W7     | Q7x2w7 uncultured   |
| 172 | 7 | 3.9 | 296 | 2 | Q885P8      | Q885p8 pseudomonas  | 245 | 7 | 3.9 | 435 | 2 | Q7U8R1     | Q7u8r1 synecococc   |
| 173 | 7 | 3.9 | 299 | 2 | Q92RQ6      | Q92rq6 rhizobium m  | 246 | 7 | 3.9 | 437 | 2 | Q7V2P6     | Q7v2p6 prochloroco  |
| 174 | 7 | 3.9 | 300 | 2 | Q72X93      | Q72x93 bacillus ce  | 247 | 7 | 3.9 | 438 | 2 | Q7V5W9     | Q7v5w9 prochloroco  |
| 175 | 7 | 3.9 | 300 | 2 | AAS44385    | Aas44385 bacillus   | 248 | 7 | 3.9 | 442 | 2 | Q8FRB8     | Q8frb8 corynebacte  |
| 176 | 7 | 3.9 | 306 | 2 | Q8W154      | Q8w154 brassica ol  | 249 | 7 | 3.9 | 443 | 2 | Q9XJC6     | Q9xjc6 streptococc  |
| 177 | 7 | 3.9 | 308 | 2 | Q8UHU1      | Q8uhj1 agrobacteri  | 250 | 7 | 3.9 | 443 | 2 | Q8SDJ6     | Q8sdj6 streptococc  |

|     |   |     |     |   |            |                     |     |   |     |      |   |            |                     |
|-----|---|-----|-----|---|------------|---------------------|-----|---|-----|------|---|------------|---------------------|
| 251 | 7 | 3.9 | 444 | 1 | TIG_RHOCA  | O68129 rhodobacter  | 324 | 7 | 3.9 | 923  | 2 | AAS65115   | Aas65115 drosophil  |
| 252 | 7 | 3.9 | 445 | 2 | Q7V0C6     | Q7v0c6 prochloroco  | 325 | 7 | 3.9 | 943  | 2 | Q8X0B4     | Q8x0b4 neurospora   |
| 253 | 7 | 3.9 | 445 | 2 | Q7VAN7     | Q7van7 prochloroco  | 326 | 7 | 3.9 | 961  | 2 | Q9F4A7     | Q9f4a7 clostridium  |
| 254 | 7 | 3.9 | 455 | 1 | NNT_YEAST  | P14743 saccharomyc  | 327 | 7 | 3.9 | 965  | 2 | Q8WZP6     | Q8wzp6 neurospora   |
| 255 | 7 | 3.9 | 458 | 2 | Q8F3N3     | Q8f3n3 leptospira   | 328 | 7 | 3.9 | 1014 | 2 | Q9RA53     | Q9ra53 thermus the  |
| 256 | 7 | 3.9 | 462 | 2 | Q9SMN8     | Q9smn8 arabidopsis  | 329 | 7 | 3.9 | 1035 | 2 | Q89ZA7     | Q89za7 bacteroides  |
| 257 | 7 | 3.9 | 462 | 2 | Q8ZV00     | Q8zv00 nitrosomona  | 330 | 7 | 3.9 | 1047 | 2 | Q9P412     | Q9p412 erysiphe gr  |
| 258 | 7 | 3.9 | 464 | 2 | Q8HDB0     | Q8hdb0 chlorogoni   | 331 | 7 | 3.9 | 1048 | 2 | Q05884     | O05884 mycobacteri  |
| 259 | 7 | 3.9 | 464 | 2 | Q8HDB3     | Q8hdb3 chlorogoni   | 332 | 7 | 3.9 | 1048 | 2 | Q7TWX3     | Q7twx3 mycobacteri  |
| 260 | 7 | 3.9 | 464 | 2 | Q8HDC2     | Q8hdc2 chlamydomon  | 333 | 7 | 3.9 | 1065 | 2 | Q8VJ44     | Q8vj44 mycobacteri  |
| 261 | 7 | 3.9 | 464 | 2 | Q6V4H8     | Q6v4h8 gastrosteu   | 334 | 7 | 3.9 | 1072 | 1 | HSEB_RAT   | P23897 rattus norv  |
| 262 | 7 | 3.9 | 464 | 2 | QAQ62978   | Aaq62978 gastrosteu | 335 | 7 | 3.9 | 1072 | 2 | Q9VI26     | Q9vi26 drosophila   |
| 263 | 7 | 3.9 | 466 | 2 | Q8R064     | Q8r064 mus musculu  | 336 | 7 | 3.9 | 1072 | 2 | Q77690     | Q77690 bos taurus   |
| 264 | 7 | 3.9 | 484 | 2 | Q7XJW3     | Q7xjw3 oryza sativ  | 337 | 7 | 3.9 | 1072 | 2 | AAF54118   | Aaf54118 drosophil  |
| 265 | 7 | 3.9 | 487 | 2 | Q750Q8     | Q750q8 ashbya goss  | 338 | 7 | 3.9 | 1073 | 1 | HSEB_HUMAN | P25092 homo sapien  |
| 266 | 7 | 3.9 | 487 | 2 | QAAS54375  | Aas54375 ashbya go  | 339 | 7 | 3.9 | 1073 | 1 | HSEB_PIG   | P55204 sus scrofa   |
| 267 | 7 | 3.9 | 490 | 2 | Q9SUM9     | Q9sum9 arabidopsis  | 340 | 7 | 3.9 | 1073 | 2 | AAB19934   | Aab19934 homo sapi  |
| 268 | 7 | 3.9 | 491 | 2 | Q84PW3     | Q84pw3 oryza sativ  | 341 | 7 | 3.9 | 1091 | 2 | Q7KX88     | Q7kx88 drosophila   |
| 269 | 7 | 3.9 | 492 | 2 | BAC99756   | Bac99756 oryza sat  | 342 | 7 | 3.9 | 1091 | 2 | AA113360   | AA113360 drosophil  |
| 270 | 7 | 3.9 | 492 | 2 | Q7ADE6     | Q7ade6 escherichia  | 343 | 7 | 3.9 | 1121 | 2 | Q8K298     | Q8k298 mus musculu  |
| 271 | 7 | 3.9 | 492 | 2 | Q7UCI7     | Q7uci7 shigella fl  | 344 | 7 | 3.9 | 1143 | 2 | Q948Y6     | Q948y6 cryptococc   |
| 272 | 7 | 3.9 | 492 | 2 | Q83L51     | Q83l51 shigella fl  | 345 | 7 | 3.9 | 1182 | 2 | Q8J121     | Q8j121 streptococc  |
| 273 | 7 | 3.9 | 492 | 2 | Q8FH01     | Q8fh01 escherichia  | 346 | 7 | 3.9 | 1184 | 2 | Q72112     | Q72112 thermus the  |
| 274 | 7 | 3.9 | 492 | 2 | Q8XDY8     | Q8xdy8 escherichia  | 347 | 7 | 3.9 | 1184 | 2 | AAS80601   | Aas80601 thermus t  |
| 275 | 7 | 3.9 | 496 | 2 | Q98EM7     | Q98em7 rhizobium l  | 348 | 7 | 3.9 | 1188 | 1 | YEC5_YEAST | P19991 saccharomyc  |
| 276 | 7 | 3.9 | 516 | 2 | Q8ZJV9     | Q8zjv9 salmonella   | 349 | 7 | 3.9 | 1215 | 2 | Q7PRV7     | Q7prv7 anopheles g  |
| 277 | 7 | 3.9 | 520 | 2 | Q8H965     | Q8h965 arabidopsis  | 350 | 7 | 3.9 | 1225 | 2 | Q7PRK2     | Q7prk2 anopheles g  |
| 278 | 7 | 3.9 | 522 | 2 | Q80Y61     | Q80y61 mus musculu  | 351 | 7 | 3.9 | 1240 | 2 | Q8R3C5     | Q8r3c5 mus musculu  |
| 279 | 7 | 3.9 | 530 | 2 | Q8VH24     | Q8vh24 cavia porce  | 352 | 7 | 3.9 | 1244 | 2 | Q94F87     | Q94f87 arabidopsis  |
| 280 | 7 | 3.9 | 541 | 2 | Q7W051     | Q7w051 bordetella   | 353 | 7 | 3.9 | 1266 | 2 | Q7KV92     | Q7kv92 drosophila   |
| 281 | 7 | 3.9 | 541 | 2 | Q7W300     | Q7w300 bordetella   | 354 | 7 | 3.9 | 1266 | 2 | AAS64939   | Aas64939 drosophil  |
| 282 | 7 | 3.9 | 551 | 2 | Q7WF70     | Q7wf70 bordetella   | 355 | 7 | 3.9 | 1275 | 2 | Q8K600     | Q8k600 streptococc  |
| 283 | 7 | 3.9 | 551 | 2 | Q89UD8     | Q89ud8 bradyrhizob  | 356 | 7 | 3.9 | 1275 | 2 | Q8NZK6     | Q8nzk6 streptococc  |
| 284 | 7 | 3.9 | 555 | 2 | Q9CCK4     | Q9cck4 mycobacteri  | 357 | 7 | 3.9 | 1275 | 2 | Q99YA0     | Q99ya0 streptococc  |
| 285 | 7 | 3.9 | 556 | 2 | Q9VMZ1     | Q9vmz1 drosophila   | 358 | 7 | 3.9 | 1277 | 2 | Q879K2     | Q879k2 streptococc  |
| 286 | 7 | 3.9 | 558 | 2 | Q9MLB2     | Q9mlb2 arabidopsis  | 359 | 7 | 3.9 | 1316 | 2 | Q9US04     | Q9us04 schizosacch  |
| 287 | 7 | 3.9 | 570 | 2 | Q9C9B0     | Q9c9b0 arabidopsis  | 360 | 7 | 3.9 | 1317 | 2 | Q6ZV29     | Q6zv29 homo sapien  |
| 288 | 7 | 3.9 | 578 | 2 | Q92WG7     | Q92wg7 rhizobium m  | 361 | 7 | 3.9 | 1317 | 2 | BAC96036   | Bac96036 homo sapi  |
| 289 | 7 | 3.9 | 579 | 2 | Q7SDB0     | Q7sdb0 neurospora   | 362 | 7 | 3.9 | 1323 | 2 | Q7TQD6     | Q7tqd6 mus musculu  |
| 290 | 7 | 3.9 | 593 | 2 | Q7RJR6     | Q7rjr6 leptospira   | 363 | 7 | 3.9 | 1323 | 2 | AAS66999   | Aas66999 mus muscu  |
| 291 | 7 | 3.9 | 593 | 2 | Q8F470     | Q8f470 leptospira   | 364 | 7 | 3.9 | 1326 | 2 | Q8BIY7     | Q8biy7 mus musculu  |
| 292 | 7 | 3.9 | 593 | 2 | QAAS70338  | Aas70338 leptospir  | 365 | 7 | 3.9 | 1327 | 2 | Q60859     | Q60859 homo sapien  |
| 293 | 7 | 3.9 | 613 | 2 | Q8K2N0     | Q8k2n0 mus musculu  | 366 | 7 | 3.9 | 1327 | 2 | Q86W58     | Q86w58 homo sapien  |
| 294 | 7 | 3.9 | 630 | 2 | Q8TAY5     | Q8tay5 homo sapien  | 367 | 7 | 3.9 | 1327 | 2 | Q8R114     | Q8r114 mus musculu  |
| 295 | 7 | 3.9 | 667 | 1 | M121_DROME | P53624 drosophila   | 368 | 7 | 3.9 | 1332 | 2 | Q8IY17     | Q8iy17 homo sapien  |
| 296 | 7 | 3.9 | 671 | 2 | Q8ALR4     | Q8alr4 bacteroides  | 369 | 7 | 3.9 | 1342 | 2 | Q6LDE9     | Q6lde9 caenorhabdi  |
| 297 | 7 | 3.9 | 672 | 2 | Q25586     | Q25586 onchocerca   | 370 | 7 | 3.9 | 1342 | 2 | AAP68922   | Aap68922 caenorhab  |
| 298 | 7 | 3.9 | 678 | 1 | AG12_DROME | Q8x3h9 drosophila   | 371 | 7 | 3.9 | 1353 | 2 | Q6LDF0     | Q6ldf0 caenorhabdi  |
| 299 | 7 | 3.9 | 682 | 2 | Q8K3H9     | Q8k3h9 rattus norv  | 372 | 7 | 3.9 | 1353 | 2 | AAP68923   | Aap68923 caenorhab  |
| 300 | 7 | 3.9 | 688 | 1 | SLP2_CLOTM | Q68853 clostridium  | 373 | 7 | 3.9 | 1357 | 2 | Q8THC8     | Q8thc8 methanosarc  |
| 301 | 7 | 3.9 | 690 | 2 | Q6P3F9     | Q6p3f9 mus musculu  | 374 | 7 | 3.9 | 1371 | 1 | YOLA_CAEEL | Q02331 caenorhabdi  |
| 302 | 7 | 3.9 | 690 | 2 | AAH64003   | Aah64003 mus muscu  | 375 | 7 | 3.9 | 1402 | 2 | Q7S8J1     | Q7s8j1 neurospora   |
| 303 | 7 | 3.9 | 702 | 2 | Q96N75     | Q96n75 homo sapien  | 376 | 7 | 3.9 | 1421 | 2 | Q7RUVO     | Q7ruvo neurospora   |
| 304 | 7 | 3.9 | 734 | 1 | PSAB_GUTH  | Q78507 guillardia   | 377 | 7 | 3.9 | 1425 | 1 | SWS_DROME  | Q9u969 drosophila   |
| 305 | 7 | 3.9 | 741 | 2 | Q6FRD4     | Q6frd4 candida gla  | 378 | 7 | 3.9 | 1441 | 2 | Q6CFC18    | Q6cf18 yarrowia li  |
| 306 | 7 | 3.9 | 754 | 2 | Q8JZQ6     | Q8jzq6 mus musculu  | 379 | 7 | 3.9 | 1441 | 2 | Q6CWC2     | Q6cwc2 kluyveromyc  |
| 307 | 7 | 3.9 | 756 | 2 | Q92FF3     | Q92ff3 listeria in  | 380 | 7 | 3.9 | 1487 | 2 | Q9KJ1      | Q9kjl1 vibrio chol  |
| 308 | 7 | 3.9 | 756 | 2 | Q8YAL3     | Q8yal3 listeria mo  | 381 | 7 | 3.9 | 1522 | 2 | Q756Z0     | Q756z0 ashbya goss  |
| 309 | 7 | 3.9 | 757 | 2 | Q75PE5     | Q75fe5 leptospira   | 382 | 7 | 3.9 | 1522 | 2 | AAS52807   | Aas52807 ashbya go  |
| 310 | 7 | 3.9 | 757 | 2 | Q8EX87     | Q8ex87 leptospira   | 383 | 7 | 3.9 | 1593 | 2 | Q8FY73     | Q8fy73 bruceella su |
| 311 | 7 | 3.9 | 757 | 2 | QAAS72270  | Aas72270 leptospir  | 384 | 7 | 3.9 | 1679 | 1 | YMF9_YEAST | Q04958 saccharomyc  |
| 312 | 7 | 3.9 | 758 | 2 | Q82U03     | Q82u03 nitrosomona  | 385 | 7 | 3.9 | 1728 | 2 | Q6FKJ21    | Q6fkj1 candida gla  |
| 313 | 7 | 3.9 | 760 | 2 | Q724V0     | Q724v0 listeria mo  | 386 | 7 | 3.9 | 1805 | 2 | Q8II22     | Q8i122 plasmodium   |
| 314 | 7 | 3.9 | 760 | 2 | QAAT02911  | Aat02911 listeria   | 387 | 7 | 3.9 | 1955 | 2 | Q8W078     | Q8w078 oryza sativ  |
| 315 | 7 | 3.9 | 778 | 2 | Q7NUI3     | Q7nu13 chromobacte  | 388 | 7 | 3.9 | 2003 | 2 | Q7LIY70    | Q7liy70 listeria mo |
| 316 | 7 | 3.9 | 790 | 2 | Q6DCQ6     | Q6dcq6 xenopus lae  | 389 | 7 | 3.9 | 2003 | 2 | AAT04744   | Aat04744 listeria   |
| 317 | 7 | 3.9 | 840 | 2 | Q98104     | Q98104 gallid heip  | 390 | 7 | 3.9 | 2105 | 2 | Q7RSY7     | Q7rsy7 giardia lam  |
| 318 | 7 | 3.9 | 848 | 2 | Q8BLB3     | Q8blb3 mus musculu  | 391 | 7 | 3.9 | 2186 | 2 | Q7QZ29     | Q7qz29 giardia lam  |
| 319 | 7 | 3.9 | 867 | 1 | EF2_BLAHO  | Q17152 blastocysti  | 392 | 7 | 3.9 | 2447 | 2 | Q6SF13     | Q6sf13 uncultured   |
| 320 | 7 | 3.9 | 872 | 2 | Q7Z9L0     | Q7z9l0 emericella   | 393 | 7 | 3.9 | 2447 | 2 | AAR38409   | Aar38409 unculture  |
| 321 | 7 | 3.9 | 877 | 2 | Q95LX7     | Q95lx7 macaca fasc  | 394 | 7 | 3.9 | 2902 | 2 | Q7WTF5     | Q7wtf5 streptomyce  |
| 322 | 7 | 3.9 | 905 | 2 | Q8SVN8     | Q8svn8 encephalito  | 395 | 7 | 3.9 | 8243 | 2 | Q96554     | Q96554 cryptospori  |
| 323 | 7 | 3.9 | 923 | 2 | Q7KX9      | Q7kx9 drosophila    | 396 | 7 | 3.9 | 25   | 2 | Q16498     | Q16498 homo sapien  |

|     |   |     |    |   |            |                    |     |   |     |     |   |            |                    |
|-----|---|-----|----|---|------------|--------------------|-----|---|-----|-----|---|------------|--------------------|
| 397 | 6 | 3.3 | 42 | 2 | Q6X9E1     | Q6x9e1 bacillus ce | 470 | 6 | 3.3 | 91  | 2 | Q50344     | Q50344 mycoplasma  |
| 398 | 6 | 3.3 | 42 | 2 | AP83796    | AP83796 bacillus   | 471 | 6 | 3.3 | 91  | 2 | BAC83710   | BAC83710 oryza sat |
| 399 | 6 | 3.3 | 42 | 2 | O50670     | O50670 borrelia bu | 472 | 6 | 3.3 | 92  | 2 | Q9UIG9     | Q9uig9 homo sapien |
| 400 | 6 | 3.3 | 54 | 2 | Q6K3V1     | Q6k3v1 oryza sativ | 473 | 6 | 3.3 | 92  | 2 | Q8U975     | Q8u975 bacterioph  |
| 401 | 6 | 3.3 | 54 | 2 | BAD19320   | BAD19320 oryza sat | 474 | 6 | 3.3 | 92  | 2 | AAQ81545   | AAQ81545 bacteriop |
| 402 | 6 | 3.3 | 56 | 2 | Q8R6T6     | Q8r6t6 thermoanaer | 475 | 6 | 3.3 | 93  | 2 | Q9ZG47     | Q9zg47 chlamydia t |
| 403 | 6 | 3.3 | 56 | 2 | Q8VJH5     | Q8vjh5 mycobacteri | 476 | 6 | 3.3 | 95  | 2 | Q9JR20     | Q9jr20 neisseria m |
| 404 | 6 | 3.3 | 59 | 2 | Q6EQW3     | Q6eqw3 oryza sativ | 477 | 6 | 3.3 | 95  | 2 | Q7DDJ8     | Q7ddj8 neisseria m |
| 405 | 6 | 3.3 | 63 | 2 | Q8E148     | Q8e148 dictyosteli | 478 | 6 | 3.3 | 96  | 2 | Q93V76     | Q93v76 oryza sativ |
| 406 | 6 | 3.3 | 64 | 2 | Q6LXL2     | Q6lxl2 methanococc | 479 | 6 | 3.3 | 96  | 2 | Q94JCI     | Q94jc1 oryza sativ |
| 407 | 6 | 3.3 | 64 | 2 | Q25245     | Q25245 lucilia cup | 480 | 6 | 3.3 | 97  | 2 | Q84TR9     | Q84tr9 oryza sativ |
| 408 | 6 | 3.3 | 64 | 2 | CAF30894   | CAF30894 methanoco | 481 | 6 | 3.3 | 97  | 2 | Q8LPZ6     | Q8lpz6 oryza sativ |
| 409 | 6 | 3.3 | 67 | 2 | Q9VE92     | Q9ve92 aeropyrum p | 482 | 6 | 3.3 | 97  | 2 | Q6EUG7     | Q6eug7 oryza sativ |
| 410 | 6 | 3.3 | 68 | 2 | Q6RJR3     | Q6rjr3 populus tom | 483 | 6 | 3.3 | 97  | 2 | Q92HX4     | Q92hx4 rickettsia  |
| 411 | 6 | 3.3 | 68 | 2 | Q8YAI0     | Q8yai0 listeria mo | 484 | 6 | 3.3 | 98  | 2 | Q7F1V3     | Q7f1v3 oryza sativ |
| 412 | 6 | 3.3 | 68 | 2 | AAR90095   | AAR90095 populus t | 485 | 6 | 3.3 | 98  | 2 | BAD02985   | BAD02985 oryza sat |
| 413 | 6 | 3.3 | 69 | 2 | Q75KP5     | Q75kp5 oryza sativ | 486 | 6 | 3.3 | 99  | 2 | Q8HIU5     | Q8hiu5 monosiga br |
| 414 | 6 | 3.3 | 69 | 2 | Q8ZAC0     | Q8zac0 yersinia pe | 487 | 6 | 3.3 | 100 | 2 | Q24920     | Q24920 echinometra |
| 415 | 6 | 3.3 | 69 | 2 | AAR87278   | AAR87278 oryza sat | 488 | 6 | 3.3 | 100 | 2 | Q24922     | Q24922 echinometra |
| 416 | 6 | 3.3 | 69 | 2 | AA563326   | AA563326 yersinia  | 489 | 6 | 3.3 | 100 | 2 | Q8VJW4     | Q8vjw4 mycobacteri |
| 417 | 6 | 3.3 | 70 | 2 | Q8SME7     | Q8sme7 caenorhabdi | 490 | 6 | 3.3 | 100 | 2 | Q6LJM7     | Q6lum7 photobacter |
| 418 | 6 | 3.3 | 72 | 2 | Q8XJ49     | Q8xj49 brucella me | 491 | 6 | 3.3 | 100 | 2 | CAG18998   | CAG18998 photobact |
| 419 | 6 | 3.3 | 74 | 2 | Q23312     | Q23312 bacillus th | 492 | 6 | 3.3 | 101 | 2 | Q6K2E3     | Q6k2e3 oryza sativ |
| 420 | 6 | 3.3 | 74 | 2 | Q74BP2     | Q74bp2 geobacter s | 493 | 6 | 3.3 | 101 | 2 | Q6Z6D1     | Q6z6d1 oryza sativ |
| 421 | 6 | 3.3 | 74 | 2 | AAR35375   | AAR35375 geobacter | 494 | 6 | 3.3 | 101 | 2 | Q9LGG6     | Q9lgg6 oryza sativ |
| 422 | 6 | 3.3 | 75 | 2 | Q74VT7     | Q74vt7 yersinia pe | 495 | 6 | 3.3 | 101 | 2 | Q94LA6     | Q94la6 arabidopsis |
| 423 | 6 | 3.3 | 75 | 2 | Q8FYQ2     | Q8fyq2 brucella su | 496 | 6 | 3.3 | 101 | 2 | Q8W0J6     | Q8w0j6 oryza sativ |
| 424 | 6 | 3.3 | 75 | 2 | AA558616   | AA558616 yersinia  | 497 | 6 | 3.3 | 101 | 2 | Q7V729     | Q7v729 prochloroco |
| 425 | 6 | 3.3 | 77 | 2 | Q96WD1     | Q96wd1 stilbella b | 498 | 6 | 3.3 | 101 | 2 | AA549074   | AA549074 arabidops |
| 426 | 6 | 3.3 | 77 | 2 | Q9KY38     | Q9ky38 streptomyce | 499 | 6 | 3.3 | 101 | 2 | BAD16015   | BAD16015 oryza sat |
| 427 | 6 | 3.3 | 77 | 2 | Q724R2     | Q724r2 listeria mo | 500 | 6 | 3.3 | 102 | 1 | RS10_MYCPE | Q8eub2 mycoplasma  |
| 428 | 6 | 3.3 | 77 | 2 | Q6Q1S0     | Q6q1s0 human coron | 501 | 6 | 3.3 | 102 | 2 | Q24918     | Q24918 echinometra |
| 429 | 6 | 3.3 | 77 | 2 | Q6R1L5     | Q6r1l5 human group | 502 | 6 | 3.3 | 102 | 2 | Q24919     | Q24919 echinometra |
| 430 | 6 | 3.3 | 77 | 2 | AA558179   | AA558179 human cor | 503 | 6 | 3.3 | 103 | 2 | Q19665     | Q19665 caenorhabdi |
| 431 | 6 | 3.3 | 77 | 2 | AAS89769   | AAS89769 human gro | 504 | 6 | 3.3 | 103 | 2 | Q7QX69     | Q7qx69 giardia lam |
| 432 | 6 | 3.3 | 77 | 2 | AAT02949   | AAT02949 listeria  | 505 | 6 | 3.3 | 103 | 2 | Q6K7M8     | Q6k7m8 oryza sativ |
| 433 | 6 | 3.3 | 78 | 2 | Q6SK01     | Q6sk01 arthrobacte | 506 | 6 | 3.3 | 103 | 2 | Q47833     | Q47833 enterococcu |
| 434 | 6 | 3.3 | 78 | 2 | AA520171   | AA520171 arthrobac | 507 | 6 | 3.3 | 103 | 2 | Q841Y2     | Q841y2 campylobact |
| 435 | 6 | 3.3 | 79 | 2 | Q6K217     | Q6k217 oryza sativ | 508 | 6 | 3.3 | 103 | 2 | BAD19481   | BAD19481 oryza sat |
| 436 | 6 | 3.3 | 79 | 2 | Q6EP17     | Q6ep17 oryza sativ | 509 | 6 | 3.3 | 104 | 1 | SUGE_PROVU | P20928 proteus vul |
| 437 | 6 | 3.3 | 79 | 2 | Q6QOM7     | Q6qom7 uncultured  | 510 | 6 | 3.3 | 104 | 2 | Q24923     | Q24923 echinometra |
| 438 | 6 | 3.3 | 79 | 2 | Q91LC1     | Q91lc1 white spot  | 511 | 6 | 3.3 | 104 | 2 | Q8HEI2     | Q8he12 varroa dest |
| 439 | 6 | 3.3 | 79 | 2 | AA591616   | AA591616 unculture | 512 | 6 | 3.3 | 105 | 1 | YE82_HAEIN | P44210 haemophilus |
| 440 | 6 | 3.3 | 80 | 1 | C553_DRSDN | P31330 desulfovibr | 513 | 6 | 3.3 | 105 | 2 | Q6R6B0     | Q6r6b0 vibriophaga |
| 441 | 6 | 3.3 | 82 | 2 | Q6VWE3     | Q6vwp3 oryza sativ | 514 | 6 | 3.3 | 105 | 2 | Q6RCE9     | Q6rce9 vibriophaga |
| 442 | 6 | 3.3 | 82 | 2 | Q6Z0J9     | Q6z0j9 oryza sativ | 515 | 6 | 3.3 | 105 | 2 | Q7EZU9     | Q7ezu9 oryza sativ |
| 443 | 6 | 3.3 | 82 | 2 | Q7CMF2     | Q7cmf2 bacillus an | 516 | 6 | 3.3 | 105 | 2 | BAC98564   | BAC98564 oryza sat |
| 444 | 6 | 3.3 | 82 | 2 | Q92RU8     | Q92ru8 rhizobium m | 517 | 6 | 3.3 | 105 | 2 | AAR92069   | AAR92069 vibrioph  |
| 445 | 6 | 3.3 | 82 | 2 | Q9X358     | Q9x358 bacillus an | 518 | 6 | 3.3 | 105 | 2 | AA97636    | AA97636 vibrioph   |
| 446 | 6 | 3.3 | 82 | 2 | BAD05648   | BAD05648 oryza sat | 519 | 6 | 3.3 | 106 | 2 | Q6K411     | Q6k411 echinometra |
| 447 | 6 | 3.3 | 82 | 2 | BAD11620   | BAD11620 oryza sat | 520 | 6 | 3.3 | 106 | 2 | Q6K411     | Q6k411 oryza sativ |
| 448 | 6 | 3.3 | 82 | 2 | BAD13256   | BAD13256 oryza sat | 521 | 6 | 3.3 | 106 | 2 | Q6Z5S2     | Q6z5s2 oryza sativ |
| 449 | 6 | 3.3 | 82 | 2 | AAT28860   | AAT28860 bacillus  | 522 | 6 | 3.3 | 106 | 2 | Q7Q760     | Q7q760 oryza sativ |
| 450 | 6 | 3.3 | 84 | 2 | Q8MTR3     | Q8mtr3 drosophila  | 523 | 6 | 3.3 | 106 | 2 | Q7XG2      | Q7xg2 oryza sativ  |
| 451 | 6 | 3.3 | 84 | 2 | Q6YZ38     | Q6yz38 oryza sativ | 524 | 6 | 3.3 | 106 | 2 | Q8LJU1     | Q8lju1 oryza sativ |
| 452 | 6 | 3.3 | 84 | 2 | Q6Z1X0     | Q6zix0 oryza sativ | 525 | 6 | 3.3 | 106 | 2 | Q7N5Z3     | Q7n5z3 photorhabdu |
| 453 | 6 | 3.3 | 84 | 2 | BAC83240   | BAC83240 oryza sat | 526 | 6 | 3.3 | 106 | 2 | Q41981     | Q41981 murid herp  |
| 454 | 6 | 3.3 | 84 | 2 | BAC92560   | BAC92560 oryza sat | 527 | 6 | 3.3 | 106 | 2 | Q98TP6     | Q98tp6 platictheys |
| 455 | 6 | 3.3 | 85 | 2 | Q8A0I4     | Q8a0i4 bacteroides | 528 | 6 | 3.3 | 106 | 2 | BAC83708   | BAC83708 oryza sat |
| 456 | 6 | 3.3 | 86 | 2 | Q8PEF8     | Q8pef8 xanthomonas | 529 | 6 | 3.3 | 107 | 2 | Q9YFZ3     | Q9yFz3 aeropyrum p |
| 457 | 6 | 3.3 | 87 | 1 | ACPY_STRAW | Q93hc3 streptomyce | 530 | 6 | 3.3 | 107 | 2 | Q8HCK7     | Q8hck7 varroa dest |
| 458 | 6 | 3.3 | 87 | 2 | Q8XWU4     | Q8xw4 oryza sativ  | 531 | 6 | 3.3 | 107 | 2 | Q6ZKV6     | Q6zkv6 oryza sativ |
| 459 | 6 | 3.3 | 87 | 2 | Q8XWZ8     | Q8xwz8 ralatonia s | 532 | 6 | 3.3 | 107 | 2 | Q8F400     | Q8f400 leptospira  |
| 460 | 6 | 3.3 | 87 | 2 | BAC98743   | BAC98743 oryza sat | 533 | 6 | 3.3 | 107 | 2 | Q9PLR9     | Q9plr9 chlamydia m |
| 461 | 6 | 3.3 | 88 | 2 | Q9XEY8     | Q9xeY8 brassica ca | 534 | 6 | 3.3 | 107 | 2 | Q8CF44     | Q8cf44 mus musculu |
| 462 | 6 | 3.3 | 88 | 2 | Q9A524     | Q9a524 oryza sativ | 535 | 6 | 3.3 | 107 | 2 | BAD05192   | BAD05192 oryza sat |
| 463 | 6 | 3.3 | 90 | 2 | Q8SRH5     | Q8srh5 encephalito | 536 | 6 | 3.3 | 108 | 2 | Q9SU51     | Q9suy1 arabidopsis |
| 464 | 6 | 3.3 | 90 | 2 | Q8MI48     | Q8mi48 bos taurus  | 537 | 6 | 3.3 | 108 | 2 | Q72RQ4     | Q72rq4 leptospira  |
| 465 | 6 | 3.3 | 90 | 2 | Q8GS65     | Q8gs65 oryza sativ | 538 | 6 | 3.3 | 108 | 2 | Q7VGF7     | Q7vgf7 helicobacte |
| 466 | 6 | 3.3 | 90 | 2 | Q6EPA7     | Q6epa7 oryza sativ | 539 | 6 | 3.3 | 108 | 2 | Q9PJU3     | Q9pju3 chlamydia m |
| 467 | 6 | 3.3 | 91 | 2 | Q7VE97     | Q7vr97 bos taurus  | 540 | 6 | 3.3 | 108 | 2 | AA570279   | AA570279 leptospir |
| 468 | 6 | 3.3 | 91 | 2 | Q8SBL3     | Q8sbl3 bacterioph  | 541 | 6 | 3.3 | 109 | 2 | Q6XN86     | Q6xn86 rhodococcu  |
| 469 | 6 | 3.3 | 91 | 2 | Q6Z5S0     | Q6z5s0 oryza sativ | 542 | 6 | 3.3 | 109 | 2 | Q6FF43     | Q6ff43 acinetobact |

|          |             |
|----------|-------------|
| Q50344   | mycoplasma  |
| BAC83710 | oryza sat   |
| Q9uig9   | homo sapien |
| Q6u975   | bacterioph  |
| AAQ81545 | bacteriop   |
| Q9zg47   | chlamydia t |
| Q9jr20   | neisseria m |
| Q7ddj8   | neisseria m |
| Q93v76   | oryza sativ |
| Q94jc1   | oryza sativ |
| Q84tr9   | oryza sativ |
| Q8lpz6   | oryza sativ |
| Q6eug7   | oryza sativ |
| Q92hx4   | rickettsia  |
| Q7f1v3   | oryza sativ |
| BAD02985 | oryza sat   |
| Q8hiu5   | monosiga br |
| Q24920   | echinometra |
| Q24922   | echinometra |
| Q8vjw4   | mycobacteri |
| Q6lum7   | photobacter |
| CAG18998 | photobact   |
| Q6k2e3   | oryza sativ |
| Q6z6d1   | oryza sativ |
| Q9lgg6   | oryza sativ |
| Q94la6   | arabidopsis |
| Q8w0j6   | oryza sativ |
| Q7v729   | prochloroco |
| AA549074 | arabidops   |
| BAD16015 | oryza sat   |
| Q8eub2   | mycoplasma  |
| Q24918   | echinometra |
| Q24919   | echinometra |
| Q19665   | caenorhabdi |
| Q7qx69   | giardia lam |
| Q6k7m8   | oryza sativ |
| Q47833   | enterococcu |
| BAD19481 | oryza sat   |
| P20928   | proteus vul |
| Q24923   | echinometra |
| Q8he12   | varroa dest |
| P44210   | haemophilus |
| Q6r6b0   | vibriophaga |
| Q6rce9   | vibriophaga |
| Q7ezu9   | oryza sativ |
| BAC98564 | oryza sat   |
| AAR92069 | vibrioph    |
| AA97636  | vibrioph    |
| Q24921   | echinometra |
| Q6k411   | oryza sativ |
| Q6z5s2   | oryza sativ |
| Q7q760   | oryza sativ |
| Q7xg2    | oryza sativ |
| Q8lju1   | oryza sativ |
| Q7n5z3   | photorhabdu |
| Q41981   | murid herp  |
| Q98tp6   | platictheys |
| BAC83708 | oryza sat   |
| Q9yFz3   | aeropyrum p |
| Q8hck7   | varroa dest |
| Q6zkv6   | oryza sativ |
| Q8f400   | leptospira  |
| Q9plr9   | chlamydia m |
| Q8cf44   | mus musculu |
| BAD05192 | oryza sat   |
| Q9suy1   | arabidopsis |
| Q72rq4   | leptospira  |
| Q7vgf7   | helicobacte |
| Q9pju3   | chlamydia m |
| AA570279 | leptospir   |
| Q6xn86   | rhodococcu  |
| Q6ff43   | acinetobact |



|     |   |     |     |   |            |                        |     |   |     |     |   |            |                    |
|-----|---|-----|-----|---|------------|------------------------|-----|---|-----|-----|---|------------|--------------------|
| 543 | 6 | 3.3 | 109 | 2 | Q8BNU2     | Q8bnu2 mus musculus    | 616 | 6 | 3.3 | 132 | 2 | Q93MP2     | Q93mp2 lactobacill |
| 544 | 6 | 3.3 | 109 | 2 | AAP73945   | Aap73945 rhodococc     | 617 | 6 | 3.3 | 132 | 2 | AAS83421   | Aas83421 oryza sat |
| 545 | 6 | 3.3 | 110 | 2 | Q9YCL0     | Q9ycl0 aeropyrum p     | 618 | 6 | 3.3 | 133 | 2 | O6XHE9     | O6xhe9 drosophila  |
| 546 | 6 | 3.3 | 110 | 2 | Q7Y0B5     | Q7y0b5 oryza sativ     | 619 | 6 | 3.3 | 133 | 2 | O8HMP9     | O8hmf9 brachydanio |
| 547 | 6 | 3.3 | 110 | 2 | O8H4E3     | O8h4e3 oryza sativ     | 620 | 6 | 3.3 | 133 | 2 | O48629     | O48629 prunus arne |
| 548 | 6 | 3.3 | 111 | 2 | O6W1H5     | O6w1h5 rhizobium s     | 621 | 6 | 3.3 | 133 | 2 | O41863     | O41863 zea mays (m |
| 549 | 6 | 3.3 | 111 | 2 | Q8Q87393   | Q8q87393 rhizobium     | 622 | 6 | 3.3 | 133 | 2 | O8SYQ5     | O8syq5 arabidopsis |
| 550 | 6 | 3.3 | 112 | 2 | Q96E07     | Q96e07 homo sapien     | 623 | 6 | 3.3 | 133 | 2 | O69552     | O69552 human herpe |
| 551 | 6 | 3.3 | 112 | 2 | Q6Z0V4     | Q6z0v4 oryza sativ     | 624 | 6 | 3.3 | 133 | 2 | AAR10257   | Aar10257 drosophil |
| 552 | 6 | 3.3 | 112 | 2 | O8VN11     | O8vn11 lactobacill     | 625 | 6 | 3.3 | 134 | 2 | O8L8Z3     | O8l8z3 arabidopsis |
| 553 | 6 | 3.3 | 112 | 2 | O985X7     | O985x7 rhizobium l     | 626 | 6 | 3.3 | 134 | 2 | O9FRF5     | O9frf5 arabidopsis |
| 554 | 6 | 3.3 | 112 | 2 | Q92UD0     | Q92ud0 rhizobium m     | 627 | 6 | 3.3 | 134 | 2 | O945L5     | O945l5 arabidopsis |
| 555 | 6 | 3.3 | 112 | 2 | Q8R7Z6     | Q8r7z6 thermoaera      | 628 | 6 | 3.3 | 134 | 2 | O8S7G8     | O8s7g8 oryza sativ |
| 556 | 6 | 3.3 | 112 | 2 | BAD16221   | Bad16221 oryza sat     | 629 | 6 | 3.3 | 134 | 2 | O8VK63     | O8vk63 mycobacteri |
| 557 | 6 | 3.3 | 113 | 2 | Q7Z6J7     | Q7z6j7 homo sapien     | 630 | 6 | 3.3 | 134 | 2 | O9IGK5     | O9igk5 pseudorabie |
| 558 | 6 | 3.3 | 113 | 2 | O6EPB7     | O6epb7 oryza sativ     | 631 | 6 | 3.3 | 135 | 2 | O7QUR1     | O7qur1 giardia lam |
| 559 | 6 | 3.3 | 113 | 2 | O8ASZ6     | O8asz6 human immun     | 632 | 6 | 3.3 | 135 | 2 | O8FV83     | O8fv83 brucella su |
| 560 | 6 | 3.3 | 114 | 1 | UL96_HHV7J | UL96_hhv7j human herpe | 633 | 6 | 3.3 | 135 | 2 | O8YD42     | O8yd42 brucella me |
| 561 | 6 | 3.3 | 114 | 2 | Q6UD64     | Q6ud64 uncultured      | 634 | 6 | 3.3 | 136 | 1 | ATPE_PROMO | ATPE_PROMO         |
| 562 | 6 | 3.3 | 114 | 2 | Q6D023     | Q6d023 erwinia car     | 635 | 6 | 3.3 | 136 | 2 | O7XIA8     | O7xia8 oryza sativ |
| 563 | 6 | 3.3 | 114 | 2 | Q77Y56     | Q77y56 human herpe     | 636 | 6 | 3.3 | 136 | 2 | O6D9P3     | O6d9p3 erwinia car |
| 564 | 6 | 3.3 | 114 | 2 | AAC40782   | Aac40782 human her     | 637 | 6 | 3.3 | 136 | 2 | O9HTQ4     | O9htq4 pseudomonas |
| 565 | 6 | 3.3 | 114 | 2 | AAR05176   | Aar05176 unculture     | 638 | 6 | 3.3 | 136 | 2 | O6EL84     | O6el84 skunkpox vi |
| 566 | 6 | 3.3 | 115 | 2 | O6I7V4     | O6i7v4 abronia gra     | 639 | 6 | 3.3 | 137 | 1 | Y137_ADE02 | Y137_ADE02         |
| 567 | 6 | 3.3 | 117 | 2 | Q99Q41     | Q99q41 streptomyce     | 640 | 6 | 3.3 | 137 | 2 | O7ZA51     | O7za51 debaryomyce |
| 568 | 6 | 3.3 | 118 | 2 | Q7U753     | Q7u753 synechococc     | 641 | 6 | 3.3 | 137 | 2 | O8KK87     | O8kk87 proteus vul |
| 569 | 6 | 3.3 | 118 | 2 | Q98450     | Q98450 paramecium      | 642 | 6 | 3.3 | 137 | 2 | O92R42     | O92r42 rhizobium m |
| 570 | 6 | 3.3 | 119 | 1 | RK22_ANTFO | RK22_antfo anthoceros  | 643 | 6 | 3.3 | 139 | 2 | O8RZ85     | O8rz85 oryza sativ |
| 571 | 6 | 3.3 | 119 | 2 | O6YXT3     | O6yxt3 oryza sativ     | 644 | 6 | 3.3 | 139 | 2 | O74F86     | O74f86 geobacter s |
| 572 | 6 | 3.3 | 119 | 2 | BAD10559   | Bad10559 oryza sat     | 645 | 6 | 3.3 | 139 | 2 | O69039     | O69039 human herpe |
| 573 | 6 | 3.3 | 120 | 2 | Q7G6D6     | Q7g6d6 oryza sativ     | 646 | 6 | 3.3 | 139 | 2 | AAR33861   | Aar33861 geobacter |
| 574 | 6 | 3.3 | 120 | 2 | O8SAX1     | O8sax1 oryza sativ     | 647 | 6 | 3.3 | 139 | 2 | O6I1K1     | O6i1k1 drosophila  |
| 575 | 6 | 3.3 | 120 | 2 | Q8GHL2     | Q8ghl2 serratia ma     | 648 | 6 | 3.3 | 140 | 2 | O6D6B9     | O6d6b9 erwinia car |
| 576 | 6 | 3.3 | 122 | 2 | O8RG15     | O8rg15 fusobacteri     | 649 | 6 | 3.3 | 140 | 2 | P79231     | P79231 physster ca |
| 577 | 6 | 3.3 | 122 | 2 | O8R3X3     | O8r3x3 mus musculu     | 650 | 6 | 3.3 | 141 | 2 | Q75G58     | Q75g58 oryza sativ |
| 578 | 6 | 3.3 | 124 | 1 | PAND_AQUAE | Pand_aquae aquifex aeo | 651 | 6 | 3.3 | 141 | 2 | AAR87360   | Aar87360 oryza sat |
| 579 | 6 | 3.3 | 124 | 2 | Q90XQ5     | Q90xd5 ambystoma m     | 652 | 6 | 3.3 | 142 | 2 | O8GVN9     | O8gvn9 oryza sativ |
| 580 | 6 | 3.3 | 125 | 2 | O8NDC5     | O8ndc5 homo sapien     | 653 | 6 | 3.3 | 143 | 2 | O8TXD0     | O8txd0 methanopyru |
| 581 | 6 | 3.3 | 125 | 2 | Q7XPV7     | Q7xpv7 oryza sativ     | 654 | 6 | 3.3 | 143 | 2 | O944Z5     | O944z5 schizosacch |
| 582 | 6 | 3.3 | 125 | 2 | O7MZD1     | O7mzd1 photorhabdu     | 655 | 6 | 3.3 | 143 | 2 | O867U7     | O867u7 trypanosom  |
| 583 | 6 | 3.3 | 125 | 2 | O8AA30     | O8aa30 bacteroides     | 656 | 6 | 3.3 | 143 | 2 | O6K9Y6     | O6k9y6 oryza sativ |
| 584 | 6 | 3.3 | 126 | 2 | O9YAU9     | O9yau9 aeropyrum p     | 657 | 6 | 3.3 | 143 | 2 | Q7X7V1     | Q7x7v1 oryza sativ |
| 585 | 6 | 3.3 | 126 | 2 | Q7G608     | Q7g608 oryza sativ     | 658 | 6 | 3.3 | 143 | 2 | O93100     | O93100 uncultured  |
| 586 | 6 | 3.3 | 126 | 2 | O94LK5     | Q94lk5 oryza sativ     | 659 | 6 | 3.3 | 143 | 2 | O837T9     | O837t9 enterococcu |
| 587 | 6 | 3.3 | 126 | 2 | O8RV03     | O8rv03 oryza sativ     | 660 | 6 | 3.3 | 143 | 2 | O8X3X2     | O8x3x2 escherichia |
| 588 | 6 | 3.3 | 127 | 1 | FUR_SYNDP7 | O55244 synechococc     | 661 | 6 | 3.3 | 144 | 2 | O7QJ81     | O7qj81 anopheles g |
| 589 | 6 | 3.3 | 127 | 2 | O8U3X4     | O8u3x4 pyrococcus      | 662 | 6 | 3.3 | 144 | 2 | O6GU91     | O6gu91 anopheles g |
| 590 | 6 | 3.3 | 127 | 2 | O9N1M1     | O9n1m1 apis mellif     | 663 | 6 | 3.3 | 144 | 2 | O6RSB3     | O6rsb3 catostomus  |
| 591 | 6 | 3.3 | 127 | 2 | O6NCC8     | O6ncc8 rhodopsedu      | 664 | 6 | 3.3 | 144 | 2 | O6RSB5     | O6rsb5 catostomus  |
| 592 | 6 | 3.3 | 127 | 2 | O8Y3M7     | O8y3m7 listeria mo     | 665 | 6 | 3.3 | 144 | 2 | O9EXG9     | O9exg9 listeria mo |
| 593 | 6 | 3.3 | 127 | 2 | CAE25988   | Ca25988 rhodopsesu     | 666 | 6 | 3.3 | 144 | 2 | AAR88224   | Aar88224 catostomu |
| 594 | 6 | 3.3 | 128 | 2 | O6C3K7     | O6c3k7 yarrowia li     | 667 | 6 | 3.3 | 144 | 2 | AAR88226   | Aar88226 catostomu |
| 595 | 6 | 3.3 | 128 | 2 | O8M241     | O8m241 drosophila      | 668 | 6 | 3.3 | 144 | 2 | AAR88228   | Aar88228 catostomu |
| 596 | 6 | 3.3 | 128 | 2 | O8LBE4     | O8lbe4 arabidopsis     | 669 | 6 | 3.3 | 144 | 2 | AAR88230   | Aar88230 catostomu |
| 597 | 6 | 3.3 | 128 | 2 | O8LCC7     | O8lcc7 arabidopsis     | 670 | 6 | 3.3 | 144 | 2 | AAR88232   | Aar88232 catostomu |
| 598 | 6 | 3.3 | 128 | 2 | Q9SU26     | Q9su26 arabidopsis     | 671 | 6 | 3.3 | 144 | 2 | AAR88234   | Aar88234 catostomu |
| 599 | 6 | 3.3 | 128 | 2 | O94AF6     | O94af6 arabidopsis     | 672 | 6 | 3.3 | 144 | 2 | AAR88236   | Aar88236 catostomu |
| 600 | 6 | 3.3 | 129 | 2 | O6ZSP9     | O6zsp9 homo sapien     | 673 | 6 | 3.3 | 145 | 2 | O761C5     | O761c5 bacterioph  |
| 601 | 6 | 3.3 | 129 | 2 | O53702     | O53702 mycobacteri     | 674 | 6 | 3.3 | 145 | 2 | O61617     | O61617 oryza sativ |
| 602 | 6 | 3.3 | 129 | 2 | O7Q2C6     | O7q2c6 desulfovibr     | 675 | 6 | 3.3 | 145 | 2 | O8W4K4     | O8w4k4 arabidopsis |
| 603 | 6 | 3.3 | 129 | 2 | Q7U260     | Q7u260 mycobacteri     | 676 | 6 | 3.3 | 145 | 2 | O66655     | O66655 aquifex aeo |
| 604 | 6 | 3.3 | 129 | 2 | BAC86898   | Bac86898 homo sapi     | 677 | 6 | 3.3 | 145 | 2 | AAQ17970   | Aaq17970 bacteriop |
| 605 | 6 | 3.3 | 129 | 2 | AAS95805   | Aas95805 desulfovi     | 678 | 6 | 3.3 | 146 | 1 | SECE_MYCLE | SECE_MYCLE         |
| 606 | 6 | 3.3 | 131 | 1 | MCRD_METFE | P12974 methanother     | 679 | 6 | 3.3 | 146 | 2 | O9HI63     | O9hi63 thermoplasm |
| 607 | 6 | 3.3 | 131 | 2 | Q96V56     | Q96v56 pleurotous o    | 680 | 6 | 3.3 | 146 | 2 | O73SG4     | O73sg4 mycobacteri |
| 608 | 6 | 3.3 | 131 | 2 | O6Z4V7     | O6z4v7 oryza sativ     | 681 | 6 | 3.3 | 147 | 2 | O82YC1     | O82yc1 streptomyce |
| 609 | 6 | 3.3 | 131 | 2 | Q6ZBH8     | Q6zbh8 oryza sativ     | 682 | 6 | 3.3 | 147 | 2 | AA506660   | AA506660 mycobacte |
| 610 | 6 | 3.3 | 131 | 2 | BAD05532   | Bad05532 oryza sat     | 683 | 6 | 3.3 | 147 | 2 | O6IG87     | O6ig87 drosophila  |
| 611 | 6 | 3.3 | 131 | 2 | BAD09730   | Bad09730 oryza sat     | 684 | 6 | 3.3 | 149 | 2 | O6K4W5     | O6k4w5 oryza sativ |
| 612 | 6 | 3.3 | 131 | 2 | BAD10715   | Bad10715 oryza sat     | 685 | 6 | 3.3 | 149 | 2 | O94173     | O94173 oryza sativ |
| 613 | 6 | 3.3 | 131 | 2 | Q9YCG5     | Q9ycg5 aeropyrum p     | 686 | 6 | 3.3 | 149 | 2 | Q6VAC9     | Q6vac9 influenza a |
| 614 | 6 | 3.3 | 132 | 2 | Q6Q7D9     | Q6q7d9 oryza sativ     | 687 | 6 | 3.3 | 149 | 2 | AAQ63112   | AAQ63112 influenza |
| 615 | 6 | 3.3 | 132 | 2 |            |                        | 688 | 6 | 3.3 | 149 | 2 |            |                    |

|     |   |     |     |   |            |                     |     |   |     |     |   |            |                     |
|-----|---|-----|-----|---|------------|---------------------|-----|---|-----|-----|---|------------|---------------------|
| 689 | 6 | 3.3 | 149 | 2 | BAD19839   | Bad19839 oryza sat  | 762 | 6 | 3.3 | 155 | 2 | CAF23258   | Caf23258 parachlam  |
| 690 | 6 | 3.3 | 150 | 1 | SODC_XENLA | P13926 xenopus lae  | 763 | 6 | 3.3 | 156 | 1 | SOD2_MESCR | O49044 mesembryant  |
| 691 | 6 | 3.3 | 150 | 1 | SODC_XENLA | P15107 xenopus lae  | 764 | 6 | 3.3 | 156 | 2 | Q76Z40     | Q76Z40 bacterioph   |
| 692 | 6 | 3.3 | 150 | 2 | O80100     | O80100 staphylococ  | 765 | 6 | 3.3 | 156 | 2 | Q8S6K0     | Q8S6K0 oryza sativ  |
| 693 | 6 | 3.3 | 150 | 2 | Q9MBQ4     | Q9MBQ4 staphylococ  | 766 | 6 | 3.3 | 156 | 2 | Q7MW64     | Q7MW64 porphyromon  |
| 694 | 6 | 3.3 | 150 | 2 | Q59795     | Q59795 pseudomonas  | 767 | 6 | 3.3 | 156 | 2 | Q9DHH9     | Q9DHH9 yaba-like d  |
| 695 | 6 | 3.3 | 150 | 2 | Q6G7W5     | Q6G7W5 staphylococ  | 768 | 6 | 3.3 | 156 | 2 | AAQ17706   | AAQ17706 bacteriop  |
| 696 | 6 | 3.3 | 150 | 2 | Q6HD59     | Q6HD59 bacillus th  | 769 | 6 | 3.3 | 157 | 1 | RL12_CHLRE | P50884 chlamydomon  |
| 697 | 6 | 3.3 | 150 | 2 | Q6CYV3     | Q6CYV3 erwinia car  | 770 | 6 | 3.3 | 157 | 2 | Q71YB2     | Q71YB2 listeria mo  |
| 698 | 6 | 3.3 | 150 | 2 | Q7Z2V9     | Q7Z2V9 bacillus ce  | 771 | 6 | 3.3 | 157 | 2 | Q6X4M2     | Q6X4M2 oncorhynch   |
| 699 | 6 | 3.3 | 150 | 2 | Q817Q7     | Q817Q7 bacillus ce  | 772 | 6 | 3.3 | 157 | 2 | AAQ94226   | AAQ94226 oncorhynch |
| 700 | 6 | 3.3 | 150 | 2 | Q81LC4     | Q81LC4 bacillus an  | 773 | 6 | 3.3 | 157 | 2 | AAT04702   | AAT04702 listeria   |
| 701 | 6 | 3.3 | 150 | 2 | Q8NVP1     | Q8NVP1 staphylococ  | 774 | 6 | 3.3 | 158 | 1 | NUDH_CANBF | Q7VRF3 candidatus   |
| 702 | 6 | 3.3 | 150 | 2 | AA543459   | AA543459 bacillus   | 775 | 6 | 3.3 | 158 | 1 | SODC_ONCVO | P24706 onchocerca   |
| 703 | 6 | 3.3 | 150 | 2 | AAT33822   | AAT33822 bacillus   | 776 | 6 | 3.3 | 158 | 2 | Q70GL0     | Q70GL0 mytilus edu  |
| 704 | 6 | 3.3 | 151 | 1 | SODI_ORISA | P28756 oryza sativ  | 777 | 6 | 3.3 | 158 | 2 | Q3Y0A5     | Q3Y0A5 acanthochei  |
| 705 | 6 | 3.3 | 151 | 1 | SODC_PEA   | Q02610 pisum sativ  | 778 | 6 | 3.3 | 158 | 2 | Q43770     | Q43770 hordeum vul  |
| 706 | 6 | 3.3 | 151 | 1 | SODC_XIPGL | P03946 xiphias gla  | 779 | 6 | 3.3 | 158 | 2 | Q9AR76     | Q9AR76 populus tre  |
| 707 | 6 | 3.3 | 151 | 2 | Q6FK19     | Q6FK19 candida gla  | 780 | 6 | 3.3 | 158 | 2 | Q9AR77     | Q9AR77 populus tre  |
| 708 | 6 | 3.3 | 151 | 2 | Q77243     | Q77243 dictyosteli  | 781 | 6 | 3.3 | 158 | 2 | Q7P8Z9     | Q7P8Z9 rickettsia   |
| 709 | 6 | 3.3 | 151 | 2 | Q6Z426     | Q6Z426 oryza sativ  | 782 | 6 | 3.3 | 158 | 2 | Q7MSC1     | Q7MSC1 wolinnella s |
| 710 | 6 | 3.3 | 151 | 2 | Q6Z9B3     | Q6Z9B3 oryza sativ  | 783 | 6 | 3.3 | 158 | 2 | Q7MSC1     | Q7MSC1 wolinnella s |
| 711 | 6 | 3.3 | 151 | 2 | Q84ZG4     | Q84ZG4 oryza sativ  | 784 | 6 | 3.3 | 158 | 2 | CAE46443   | CAE46443 mytilus e  |
| 712 | 6 | 3.3 | 151 | 2 | Q93VE4     | Q93VE4 arabidopsis  | 785 | 6 | 3.3 | 159 | 2 | Q9HAV2     | Q9HAV2 homo sapien  |
| 713 | 6 | 3.3 | 151 | 2 | Q93VE4     | Q93VE4 arabidopsis  | 786 | 6 | 3.3 | 159 | 2 | Q9LXV9     | Q9LXV9 arabidopsis  |
| 714 | 6 | 3.3 | 151 | 2 | Q941A0     | Q941A0 arabidopsis  | 787 | 6 | 3.3 | 159 | 2 | Q9K5F8     | Q9K5F8 bacillus ce  |
| 715 | 6 | 3.3 | 151 | 2 | BAD05461   | Bad01293 oryza sat  | 788 | 6 | 3.3 | 159 | 2 | Q7CPG5     | Q7CPG5 salmonella   |
| 716 | 6 | 3.3 | 151 | 2 | BAD05461   | Bad05461 oryza sat  | 789 | 6 | 3.3 | 159 | 2 | Q9R0F1     | Q9R0F1 salmonella   |
| 717 | 6 | 3.3 | 151 | 2 | BAD13068   | Bad13068 oryza sat  | 790 | 6 | 3.3 | 160 | 2 | Q9ELZ6     | Q9ELZ6 debaromyce   |
| 718 | 6 | 3.3 | 151 | 2 | AAH70696   | AAH70696 xenopus l  | 791 | 6 | 3.3 | 160 | 2 | Q9RQW9     | Q9RQW9 plasmodium   |
| 719 | 6 | 3.3 | 152 | 2 | Q6YWI1     | Q6YWI1 oryza sativ  | 792 | 6 | 3.3 | 160 | 2 | Q9ZUP5     | Q9ZUP5 rhizobium m  |
| 720 | 6 | 3.3 | 152 | 2 | Q9ZNU4     | Q9ZNU4 cicier ariet | 793 | 6 | 3.3 | 161 | 1 | ISPF_DEIRA | Q9Y5J9 anabaena sp  |
| 721 | 6 | 3.3 | 152 | 2 | Q6EQL6     | Q6EQL6 oryza sativ  | 794 | 6 | 3.3 | 161 | 1 | SECE_MYCTU | Q9RKS6 deinococcus  |
| 722 | 6 | 3.3 | 152 | 2 | Q88MZ3     | Q88MZ3 pseudomonas  | 795 | 6 | 3.3 | 161 | 1 | Q7XT20     | P96929 mycobacteri  |
| 723 | 6 | 3.3 | 152 | 2 | Q6LDS4     | Q6LDS4 rattus norv  | 796 | 6 | 3.3 | 161 | 2 | Q7XT20     | Q7XT20 oryza sativ  |
| 724 | 6 | 3.3 | 152 | 2 | BAC81657   | BAC81657 pisum sat  | 797 | 6 | 3.3 | 161 | 2 | Q7XUX1     | Q7XUX1 oryza sativ  |
| 725 | 6 | 3.3 | 152 | 2 | AAA40996   | AAA40996 rattus no  | 798 | 6 | 3.3 | 161 | 2 | Q8YTM2     | Q8YTM2 anabaena sp  |
| 726 | 6 | 3.3 | 152 | 2 | BAC98746   | BAC98746 oryza sat  | 799 | 6 | 3.3 | 161 | 2 | Q8Y5S4     | Q8Y5S4 listeria mo  |
| 727 | 6 | 3.3 | 153 | 1 | SODC_RAT   | P07632 rattus norv  | 800 | 6 | 3.3 | 161 | 2 | Q71Y40     | Q71Y40 listeria mo  |
| 728 | 6 | 3.3 | 153 | 1 | SODC_YEAST | P00445 saccharomyc  | 801 | 6 | 3.3 | 161 | 2 | AAT04774   | AAT04774 listeria   |
| 729 | 6 | 3.3 | 153 | 2 | Q9YAJ3     | Q9YAJ3 aeropyrum p  | 802 | 6 | 3.3 | 162 | 1 | HRS2_HUMAN | Q9NWV9 homo sapien  |
| 730 | 6 | 3.3 | 153 | 2 | Q8GF96     | Q8GF96 schistosoma  | 803 | 6 | 3.3 | 162 | 2 | Q9UXI9     | Q9UXI9 sulfolobus   |
| 731 | 6 | 3.3 | 153 | 2 | Q8I807     | Q8I807 anemonia vi  | 804 | 6 | 3.3 | 162 | 2 | O81236     | O81236 arabidopsis  |
| 732 | 6 | 3.3 | 153 | 2 | Q67GK6     | Q67GK6 bos taurus   | 805 | 6 | 3.3 | 162 | 2 | Q6ZEY5     | Q6ZEY5 oryza sativ  |
| 733 | 6 | 3.3 | 153 | 2 | Q8WNN6     | Q8WNN6 canis famil  | 806 | 6 | 3.3 | 163 | 2 | Q6PEC5     | Q6PEC5 rattus norv  |
| 734 | 6 | 3.3 | 153 | 2 | Q701I0     | Q701I0 arabidopsis  | 807 | 6 | 3.3 | 163 | 2 | AAH58148   | AAH58148 rattus no  |
| 735 | 6 | 3.3 | 153 | 2 | Q9I438     | Q9I438 pseudomonas  | 808 | 6 | 3.3 | 163 | 2 | BAC83446   | BAC83446 oryza sat  |
| 736 | 6 | 3.3 | 153 | 2 | AAQ96173   | AAQ96173 bos tauru  | 809 | 6 | 3.3 | 164 | 2 | Q9BK00     | Q9BK00 giardia lam  |
| 737 | 6 | 3.3 | 153 | 2 | CAE45016   | CAE45016 arabidops  | 810 | 6 | 3.3 | 164 | 2 | Q9SVX9     | Q9SVX9 giardia lam  |
| 738 | 6 | 3.3 | 153 | 2 | AA598801   | AA598801 anemonia   | 811 | 6 | 3.3 | 164 | 2 | Q6S8F4     | Q6S8F4 musa acumin  |
| 739 | 6 | 3.3 | 154 | 2 | Q750W5     | Q750W5 ashbya gos   | 812 | 6 | 3.3 | 164 | 2 | Q9FK60     | Q9FK60 arabidopsis  |
| 740 | 6 | 3.3 | 154 | 2 | Q6FWA5     | Q6FWA5 candida gla  | 813 | 6 | 3.3 | 164 | 2 | Q6GB30     | Q6GB30 staphylococ  |
| 741 | 6 | 3.3 | 154 | 2 | Q6FWL5     | Q6FWL5 candida gla  | 814 | 6 | 3.3 | 164 | 2 | Q6GIJ0     | Q6GIJ0 staphylococ  |
| 742 | 6 | 3.3 | 154 | 2 | O24127     | O24127 nicotiana t  | 815 | 6 | 3.3 | 164 | 2 | Q93I01     | Q93I01 uncultured   |
| 743 | 6 | 3.3 | 154 | 2 | Q61YF2     | Q61YF2 pseudomonas  | 816 | 6 | 3.3 | 164 | 2 | Q9PNA4     | Q9PNA4 campylobact  |
| 744 | 6 | 3.3 | 154 | 2 | Q61YF2     | Q61YF2 oplegnathus  | 817 | 6 | 3.3 | 164 | 2 | Q99VI0     | Q99VI0 staphylococ  |
| 745 | 6 | 3.3 | 154 | 2 | Q6RUW1     | Q6RUW1 oreochromis  | 818 | 6 | 3.3 | 164 | 2 | Q7A1E7     | Q7A1E7 staphylococ  |
| 746 | 6 | 3.3 | 154 | 2 | Q8QHI0     | Q8QHI0 oncorhynch   | 819 | 6 | 3.3 | 164 | 2 | Q7A6N1     | Q7A6N1 staphylococ  |
| 747 | 6 | 3.3 | 154 | 2 | AA82969    | AA82969 oreochrom   | 820 | 6 | 3.3 | 164 | 2 | Q8R4S0     | Q8R4S0 mus musculu  |
| 748 | 6 | 3.3 | 154 | 2 | AA556399   | AA556399 saccharom  | 821 | 6 | 3.3 | 164 | 2 | AA888780   | AA888780 musa acum  |
| 749 | 6 | 3.3 | 154 | 2 | AA554315   | AA554315 ashbya go  | 822 | 6 | 3.3 | 165 | 2 | Q93HZ8     | Q93HZ8 uncultured   |
| 750 | 6 | 3.3 | 154 | 2 | AAT36615   | AAT36615 oplegnath  | 823 | 6 | 3.3 | 165 | 2 | Q93HZ9     | Q93HZ9 uncultured   |
| 751 | 6 | 3.3 | 155 | 2 | Q6RBY0     | Q6RBY0 biomphali    | 824 | 6 | 3.3 | 165 | 2 | Q9PNA6     | Q9PNA6 campylobact  |
| 752 | 6 | 3.3 | 155 | 2 | Q7YXL9     | Q7YXL9 lymanaea sta | 825 | 6 | 3.3 | 165 | 2 | Q9Y502     | Q9Y502 listeria mo  |
| 753 | 6 | 3.3 | 155 | 2 | Q8WUT8     | Q8WUT8 aplysia cal  | 826 | 6 | 3.3 | 166 | 1 | Y504_MVCLL | P54979 mycobacteri  |
| 754 | 6 | 3.3 | 155 | 2 | Q9LWK6     | Q9LWK6 arabidopsis  | 827 | 6 | 3.3 | 166 | 2 | Q14827     | Q14827 homo sapien  |
| 755 | 6 | 3.3 | 155 | 2 | Q6MDU1     | Q6MDU1 perachlamyd  | 828 | 6 | 3.3 | 166 | 2 | Q38279     | Q38279 lactococcus  |
| 756 | 6 | 3.3 | 155 | 2 | Q76DK7     | Q76DK7 xenopus lae  | 829 | 6 | 3.3 | 166 | 2 | Q6NZ29     | Q6NZ29 rhodopsu     |
| 757 | 6 | 3.3 | 155 | 2 | AAQ89649   | AAQ89649 arabidops  | 830 | 6 | 3.3 | 166 | 2 | Q7NGS0     | Q7NGS0 pseudomonas  |
| 758 | 6 | 3.3 | 155 | 2 | BAC87700   | BAC87700 xenopus l  | 831 | 6 | 3.3 | 166 | 2 | Q9HXN9     | Q9HXN9 rhodopsu     |
| 759 | 6 | 3.3 | 155 | 2 | AA93637    | AA93637 lymanaea s  | 832 | 6 | 3.3 | 167 | 2 | CAE29339   | CAE29339 rhodopsu   |
| 760 | 6 | 3.3 | 155 | 2 | AA93637    | AA93637 biomphala   | 833 | 6 | 3.3 | 167 | 2 | Q6CTH3     | Q6CTH3 kluyveromyc  |
| 761 | 6 | 3.3 | 155 | 2 | AA93637    | AA93637 biomphala   | 834 | 6 | 3.3 | 167 | 2 | Q75134     | Q75134 oryza sativ  |

|     |   |     |     |   |            |                     |     |   |     |     |   |            |                    |
|-----|---|-----|-----|---|------------|---------------------|-----|---|-----|-----|---|------------|--------------------|
| 835 | 6 | 3.3 | 167 | 2 | Q845B7     | Q845b7 tropheryma   | 908 | 6 | 3.3 | 181 | 2 | Q84R48     | Q84r48 oryza sativ |
| 836 | 6 | 3.3 | 167 | 2 | Q6N604     | Q6n604 rhodopseu    | 909 | 6 | 3.3 | 181 | 2 | Q93I35     | Q93i35 pseudomonas |
| 837 | 6 | 3.3 | 167 | 2 | Q9Z6G2     | Q9z6g2 vibrio para  | 910 | 6 | 3.3 | 181 | 2 | Q8DM9      | Q8dm9 synechococ   |
| 838 | 6 | 3.3 | 167 | 2 | Q7AR94     | Q7ar94 campylobact  | 911 | 6 | 3.3 | 181 | 2 | Q9QUC9     | Q9quc9 tt virus. o |
| 839 | 6 | 3.3 | 167 | 2 | Q8AR7213   | Aar87213 oryza sat  | 912 | 6 | 3.3 | 181 | 2 | Q9QUD1     | Q9qud1 tt virus. o |
| 840 | 6 | 3.3 | 167 | 2 | CAE828256  | CaE828256 rhodopseu | 913 | 6 | 3.3 | 181 | 2 | Q90Y11     | Q90y11 crotalus du |
| 841 | 6 | 3.3 | 168 | 2 | Q9XXD0     | Q9xxd0 caenorhabdi  | 914 | 6 | 3.3 | 181 | 2 | Q90Y12     | Q90y12 crotalus du |
| 842 | 6 | 3.3 | 168 | 2 | Q7XQD8     | Q7xqd8 oryza sativ  | 915 | 6 | 3.3 | 181 | 2 | Q90Y12     | Q90y12 crotalus du |
| 843 | 6 | 3.3 | 168 | 2 | Q8FB84     | Q8fb84 escherichia  | 916 | 6 | 3.3 | 181 | 2 | Q90Y12     | Q90y12 crotalus du |
| 844 | 6 | 3.3 | 168 | 2 | Q92I10     | Q92i10 rickettsia   | 917 | 6 | 3.3 | 181 | 2 | Q90Y12     | Q90y12 crotalus du |
| 845 | 6 | 3.3 | 169 | 1 | MINC_DEIRA | Q9rtk7 deinococcus  | 918 | 6 | 3.3 | 182 | 2 | Q7XV35     | Q7xv35 oryza sativ |
| 846 | 6 | 3.3 | 169 | 1 | SODJ_AQUAE | Q67l49 aquifex aeo  | 919 | 6 | 3.3 | 182 | 2 | Q8S1T5     | Q8s1t5 oryza sativ |
| 847 | 6 | 3.3 | 169 | 2 | Q859F9     | Q859f9 pseudomonas  | 920 | 6 | 3.3 | 182 | 2 | Q93I27     | Q93i27 pseudomonas |
| 848 | 6 | 3.3 | 169 | 2 | Q9KHA3     | Q9kha3 nostoc punc  | 921 | 6 | 3.3 | 182 | 2 | Q7VX08     | Q7vx08 bordetella  |
| 849 | 6 | 3.3 | 169 | 2 | Q8XKS3     | Q8xks3 clostridium  | 922 | 6 | 3.3 | 182 | 2 | Q97JC3     | Q97jc3 clostridium |
| 850 | 6 | 3.3 | 170 | 2 | Q6HSR1     | Q6hsr1 oryza sativ  | 923 | 6 | 3.3 | 182 | 2 | Q8BX39     | Q8bx39 mus musculu |
| 851 | 6 | 3.3 | 170 | 2 | Q8BKX8     | Q8bkx8 mus musculu  | 924 | 6 | 3.3 | 182 | 2 | Q66244     | Q66244 citrus tris |
| 852 | 6 | 3.3 | 171 | 2 | Q85021     | Q85021 photobacter  | 925 | 6 | 3.3 | 183 | 2 | Q6CUE0     | Q6cue0 kluyveromyc |
| 853 | 6 | 3.3 | 172 | 2 | Q6Z552     | Q6z552 oryza sativ  | 926 | 6 | 3.3 | 183 | 2 | Q93I28     | Q93i28 pseudomonas |
| 854 | 6 | 3.3 | 172 | 2 | Q887F8     | Q887f8 pseudomonas  | 927 | 6 | 3.3 | 183 | 2 | Q93I33     | Q93i33 pseudomonas |
| 855 | 6 | 3.3 | 172 | 2 | Q911L23    | Q911l23 citrus tris | 928 | 6 | 3.3 | 184 | 1 | DEF2_BACAN | Q81mc9 bacillus an |
| 856 | 6 | 3.3 | 172 | 2 | Q911L23    | Q911l23 citrus tris | 929 | 6 | 3.3 | 184 | 1 | DEF2_BACCR | Q819k2 bacillus ce |
| 857 | 6 | 3.3 | 173 | 2 | Q28289     | Bac99708 oryza sat  | 930 | 6 | 3.3 | 184 | 1 | DEF2_BACST | Q31410 bacillus st |
| 858 | 6 | 3.3 | 173 | 2 | Q6Z7U4     | Q6z7u4 canis famil  | 931 | 6 | 3.3 | 184 | 2 | Q96GI7     | Q96gi7 homo sapien |
| 859 | 6 | 3.3 | 173 | 2 | Q84R45     | Q84r45 oryza sativ  | 932 | 6 | 3.3 | 184 | 2 | Q7Q744     | Q7q744 anopheles g |
| 860 | 6 | 3.3 | 173 | 2 | Q9FZ17     | Q9fzi7 arabidopsis  | 933 | 6 | 3.3 | 184 | 2 | Q6RCL4     | Q6rc14 pseudomonas |
| 861 | 6 | 3.3 | 173 | 2 | Q6N247     | Q6n247 rhodopseu    | 934 | 6 | 3.3 | 184 | 2 | Q6HEJ1     | Q6hej1 bacillus th |
| 862 | 6 | 3.3 | 173 | 2 | Q89QZ7     | Q89qz7 bradyrhizob  | 935 | 6 | 3.3 | 184 | 2 | Q6HMK8     | Q6hm8 bacillus th  |
| 863 | 6 | 3.3 | 173 | 2 | Q8D332     | Q8d332 wiggiwort    | 936 | 6 | 3.3 | 184 | 2 | Q731Z2     | Q731z2 bacillus ce |
| 864 | 6 | 3.3 | 173 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 937 | 6 | 3.3 | 184 | 2 | Q7MNC5     | Q7mnc5 vibrio vuln |
| 865 | 6 | 3.3 | 173 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 938 | 6 | 3.3 | 184 | 2 | Q7UF44     | Q7uf44 rhodopirell |
| 866 | 6 | 3.3 | 173 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 939 | 6 | 3.3 | 184 | 2 | Q813X0     | Q813x0 bacillus ce |
| 867 | 6 | 3.3 | 173 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 940 | 6 | 3.3 | 184 | 2 | Q81UC5     | Q81uc5 bacillus an |
| 868 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 941 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 869 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 942 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 870 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 943 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 871 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 944 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 872 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 945 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 873 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 946 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 874 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 947 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 875 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 948 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 876 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 949 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 877 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 950 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 878 | 6 | 3.3 | 174 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 951 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 879 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 952 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 880 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 953 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 881 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 954 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 882 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 955 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 883 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 956 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 884 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 957 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 885 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 958 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 886 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 959 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 887 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 960 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 888 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 961 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 889 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 962 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 890 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 963 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 891 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 964 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 892 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 965 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 893 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 966 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 894 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 967 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 895 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 968 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 896 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 969 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 897 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 970 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 898 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 971 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 899 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 972 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 900 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 973 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 901 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 974 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 902 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 975 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 903 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 976 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 904 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 977 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 905 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 978 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 906 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 979 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |
| 907 | 6 | 3.3 | 175 | 2 | Q89QZ7     | Q89qz7 wiggiwort    | 980 | 6 | 3.3 | 184 | 2 | Q8DF19     | Q8df19 vibrio vuln |

```

981 Q8csm3 staphylococ
982 Q54225 streptomyce
983 Q8b7w5 oryza sativ
984 P41974 difofilaria
985 Q8n7l3 homo sapien
986 Q6z6t7 oryza sativ
987 Q93i34 pseudomonas
988 Q9rwj5 deinococcus
989 Bad15958 oryza sat
990 Q6bmz7 debaryomyce
991 Q749x7 geobacter s
992 Q927z6 chlamydia p
993 Q92wfw4 rhizobium m
994 Aar35987 geobacter
995 P46851 escherichia
996 Q97c34 thermoplasm
997 Q7pk04 anopheles g
998 Q9njs2 leishmania
999 Q8m165 erinaceus e
1000 Q94rn6 stilodes fu

```

## ALIGNMENTS

```

RESULT 1
Q9MC02 ID Q9MC02 PRELIMINARY; PRT; 174 AA.
AC Q9MC02;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative Cu/Zn superoxide dismutase.
GN Name=sodCIII;
OS Phage FeIs-1.
OC Viruses.
OX NCBI_TaxID=128975;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21065122; PubMed=11136448;
RA Figueroa-Bossi N., Uzau S., Maloriot D., Bossi L.;
RT "Variable assortment of prophages provides a transferable repertoire
RT of pathogenic determinants in Salmonella."
RL Mol. Microbiol. 39:260-271(2001).
DR EMBL; AF254764; AAF82484.1; -.
DR HSSP; P53636; 1EQW.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN_1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
SQ SEQUENCE 174 AA; 17995 MW; 557DE7F2B5DB76D0 CRC64;

Query Match 7.8%; Score 14; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 165 PLGGGGGARIACGVI 178
DB 160 PLGGGGGARIACGVI 173

RESULT 2
Q8ZQF7 ID Q8ZQF7 PRELIMINARY; PRT; 174 AA.
AC Q8ZQF7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative FeIs-1 prophage Cu/Zn superoxide dismutase (EC 1.15.1.1)

```

```

DE (Superoxide dismutase [Cu-Zn]).
GN OrderedLocusNames=STM0924;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AE008873; AAL19860.1; -.
DR HSSP; P53636; 1EQW.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN_1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 174 AA; 17995 MW; 557DE7F2B5DB76D0 CRC64;

```

```

Query Match 7.8%; Score 14; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGGARIACGVI 178
DB 160 PLGGGGGARIACGVI 173

RESULT 3
Q6G3C6 ID Q6G3C6 PRELIMINARY; PRT; 175 AA.
AC Q6G3C6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Superoxide dismutase.
GN Name=sodC; OrderedLocusNames=BH08570;
OS Bartonella henselae (Rochalimaea henselae).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=38323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49882 / Houston 1;
RX PubMed=15210978;
RA Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,
RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,
RA La Scola B., Holmberg M., Andersson S.G.E.;
RT "The louse-borne human pathogen Bartonella quintana is a genomic
RT derivative of the zoonotic agent Bartonella henselae."
RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

```

CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By similarity).

CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.

DR EMBL: BX897699; CFA27655.1; -.

DR InterPro: IPR001424; SOD\_CU\_ZN.

DR Pfam: PF00080; SOD\_Cu; 1.

DR ProDom: PD000469; SOD\_CU\_ZN; 1.

DR PROSITE: PS00087; SOD\_CU\_ZN\_1; 1.

DR PROSITE: PS00332; SOD\_CU\_ZN\_2; 1.

KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.

SQ SEQUENCE 175 AA; 18661 MW; 5CC93A4ADB35D9E CRC64;

Query Match 7.2%; Score 13; DB 2; Length 175;

Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 SDKPLPLGGGAR 172  
 Db 155 SDKPLPLGGGAR 167  
 |||||

RESULT 4

Q8ELY4 PRELIMINARY; PRT; 194 AA.

AC Q8ELY4;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DR 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DE Superoxide dismutase (EC 1.15.1.1).

GN OrderedLocustNames=OB3080;

OS Oceanobacillus iheyensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OX NCBI\_TaxID=182710;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HTE831;

RX MEDLINE=22220767; PubMed=12335376;

RA Takami H., Takaki Y., Uchiyama I.;

RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments."

RT Nucleic Acids Res. 30:3927-3935(2002).

CC -!- FUNCTION: Destroys radicals which are normally produced within the cells and which are toxic to biological systems (By similarity).

CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By similarity).

CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.

DR EMBL: AF004603; BAC15036.1; -.

DR HSSP: P00441; 1OZU.

DR GO: GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.

DR GO: GO:0046872; F:metal ion binding; IEA.

DR GO: GO:0016491; F:oxidoreductase activity; IEA.

DR GO: GO:0008801; P:superoxide metabolism; IEA.

DR InterPro: IPR001424; SOD\_CU\_ZN.

DR Pfam: PF00080; SOD\_Cu; 1.

DR PRINTS: PR00068; CUZNDISMTASE.

DR ProDom: PD000469; SOD\_CU\_ZN; 1.

KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.

SQ SEQUENCE 194 AA; 20816 MW; 1843B4024A3C36 CRC64;

Query Match 6.7%; Score 12; DB 2; Length 194;

Best Local Similarity 100.0%; Pred. No. 0.0015;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LPAGEHGFIHE 77  
 Db 81 LPAGEHGFIHE 92  
 |||||

RESULT 5

Q83A08 PRELIMINARY; PRT; 170 AA.

ID Q83A08

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Superoxide dismutase.

GN Name=sodC; OrderedLocustNames=CBU1822;

OS Coccidia burnetii.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;

OC Coccidiaceae; Coccidia.

OX NCBI\_TaxID=777;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Nine Mile Phase I / RSA 493;

RX MEDLINE=22608657; PubMed=12704232;

RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E., Nelson W.C., Ward N.L., Tettelin H., Daviden T.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J., Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A., Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;

RT "Complete genome sequence of the Q-fever pathogen, Coccidia burnetii."

RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).

CC -!- FUNCTION: Destroys radicals which are normally produced within the cells and which are toxic to biological systems (By similarity).

CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By similarity).

CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.

DR EMBL: AE016965; AAO91315.1; -.

DR HSSP: P53636; 1EQW.

DR TIGR: CBU1822; -.

DR GO: GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.

DR GO: GO:0046872; F:metal ion binding; IEA.

DR GO: GO:0008801; P:superoxide metabolism; IEA.

DR InterPro: IPR001424; SOD\_CU\_ZN.

DR Pfam: PF00080; SOD\_Cu; 1.

DR PRINTS: PR00068; CUZNDISMTASE.

DR ProDom: PD000469; SOD\_CU\_ZN; 1.

DR PROSITE: PS00087; SOD\_CU\_ZN\_1; UNKNOWN 1.

DR PROSITE: PS00332; SOD\_CU\_ZN\_2; UNKNOWN 1.

KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.

SQ SEQUENCE 170 AA; 17871 MW; 5CB62830C4C2DE04 CRC64;

Query Match 6.1%; Score 11; DB 2; Length 170;

Best Local Similarity 100.0%; Pred. No. 0.014;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDNYSD 161  
 Db 142 MIHAGGDNYSD 152  
 |||||

RESULT 6

SODC\_FRATU STANDARD; PRT; 185 AA.

ID SODC\_FRATU

AC Q59438;

DT 15-DEC-1998 (rel. 37, Created)

DT 15-DEC-1998 (rel. 37, Last sequence update)

DT 05-JUL-2004 (rel. 44, Last annotation update)

DE Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1).

GN Name=sodC;

OS Francisella tularensis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales;

OC Francisellaceae; Francisella.

OX NCBI\_TaxID=263;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=UVS;

RA Clairoux N., Nano F.E., Boissinot M.;

RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Destroys radicals which are normally produced within the cells and which are toxic to biological systems (By similarity).

CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).



```

DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 201 AA; 26675 MW; 75D25AB848BF8DCD CRC64;

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 201;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDNYSYD 161
DB 173 MIHAGGDNYSYD 183

RESULT 9
AAS60586 PRELIMINARY; PRT; 201 AA.
ID AAS60586;
AC AAS60586;
DT 24-MAR-2004 (TrEMBLrel. 27, Created)
DT 24-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Superoxide dismutase.
GN SODC OR YP0311.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017128; AAS60586.1; -.
SQ SEQUENCE 201 AA; 20675 MW; 75D25AB848BF8DCD CRC64;

Query Match
Best Local Similarity 100.0%; Score 11; DB 2; Length 201;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDNYSYD 161
DB 173 MIHAGGDNYSYD 183

RESULT 10
SODC_BRUAB STANDARD; PRT; 154 AA.
ID SODC_BRUAB;
AC P15453;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN Name=sodc;
OS Brucella abortus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP SEQUENCE.
RX MEDLINE=90148961; PubMed=2105741;
RA Beck B.L., Tabatabai L.B., Mayfield J.E.;
RT "A protein isolated from Brucella abortus is a Cu-Zn superoxide
dismutase."
RL Biochemistry 29:372-376(1990).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
similarity).

```

```

CC similarity).
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR PIR; A38993; A38993.
DR HSP; P53636; IEQW.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Antioxidant; Copper; Direct protein sequencing; Metal-binding;
Oxidoreductase; Periplasmic; Zinc.
FT METAL 48 48 Copper (By similarity).
FT METAL 50 50 Copper (By similarity).
FT METAL 73 73 Copper and zinc (By similarity).
FT METAL 82 82 Zinc (By similarity).
FT METAL 90 90 Zinc (By similarity).
FT METAL 93 93 Zinc (By similarity).
FT METAL 128 128 Copper (By similarity).
FT DISULFID 55 150 By similarity.
SQ SEQUENCE 154 AA; 16072 MW; 4672C31481704468 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 1; Length 154;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDKP 163
DB 130 GGDNYSDKP 138

RESULT 11
Q7MFM9 PRELIMINARY; PRT; 170 AA.
ID Q7MFM9;
AC Q7MFM9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Superoxide dismutase, Cu-Zn.
GN Name=VWA0291;
OS Vibrio vulnificus (strain Y0016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-P.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
pathogen."
RT Genome Res. 13:2577-2587(2003).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems (By similarity).
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
similarity).
CC -1- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AF005345; BAC96317.1; -.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN 1.
KW Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 170 AA; 17613 MW; 5AA2C93F1176704A CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 2; Length 170;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 165 PLGGGGARI 173
Db 156 PLGGGGARI 164

RESULT 12
Q8D454 PRELIMINARY; PRT; 170 AA.
AC Q8D454;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Cu/Zn superoxide dismutase.
GN OrderedLocusNames=VW21471;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Khee J.H., Kim S.-Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COPACITOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
CC EMBL; AE016813; AA08341.1; -.
DR HSP; P00446; 10AL.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; SOD_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN; 1; UNKNOWN 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 170 AA; 17585 MW; AAB5DE3F1766977 CRC64;

Query Match 5.0%; Score 9; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 165 PLGGGGARI 173
Db 156 PLGGGGARI 164

RESULT 13
Q9KRQ3 PRELIMINARY; PRT; 171 AA.
AC Q9KRQ3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Superoxide dismutase, Cu-Zn.
GN OrderedLocusNames=VCL1583;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406633; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,

```

```

RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COPACITOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
CC EMBL; AE004235; AAF94737.1; -.
DR PIR; A82183; A82183.
DR HSP; P00446; 1YAI.
DR TIGR; VC1583; -.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:00046872; F:metal ion binding; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; SOD_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN; 1; UNKNOWN 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 171 AA; 17697 MW; 7483250CB4266C79 CRC64;

Query Match 5.0%; Score 9; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 165 PLGGGGARI 173
Db 156 PLGGGGARI 164

RESULT 14
SODC BRUME
ID _SODC BRUME STANDARD; PRT; 174 AA.
AC P58645;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Superoxide dismutase [Cu-Zn] precursor [EC 1.15.1.1].
GN Name=sodC; OrderedLocusNames=BMEII0581, BRA0703;
OS Brucella melitensis, and
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459, 29461;
RN [1]_
RP SEQUENCE FROM N.A.
RC SPECIES=B.melitensis; STRAIN=16M / ATCC 23456 / Biotype 1;
RX PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kaparatil V., Redkar R.J., Patra G., Mujeer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Iykidis A., Resnik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
RN [2]_
RP SEQUENCE FROM N.A.
RC SPECIES=B.suis; STRAIN=1330 / Biovar 1;
RX MEDLINE=2247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kral M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,

```



```

RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RL animal and plant pathogens and symbionts.";
RC Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC -|- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -|- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Periplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE009694; AAL53823.1; -.
CC FIRM; AE014566; AAN33888.1; ALT_INIT.
CC PIR; AD3582; AD3582.
CC HSP; PS3636; IEQM.
CC TIGR; BRA0703; -.
CC InterPro; IPR001424; SOD_CU_ZN.
CC ProDom; PD000469; SOD_CU_ZN; 1.
CC PROSITE; PS00087; SOD_CU_ZN_1; 1.
CC PROSITE; PS00332; SOD_CU_ZN_2; 1.
CC Antioxidant; Complete proteome; Copper; Metal-binding; Oxidoreductase;
KW Periplasmic; Signal; Zinc.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 174 Superoxide dismutase [Cu-Zn].
FT METAL 68 68 Copper (By similarity).
FT METAL 70 70 Copper (By similarity).
FT METAL 93 93 Copper and zinc (By similarity).
FT METAL 102 102 Zinc (By similarity).
FT METAL 110 110 Zinc (By similarity).
FT METAL 113 113 Zinc (By similarity).
FT METAL 148 148 Copper (By similarity).
FT DISULFID 75 170 By similarity.
SQ SEQUENCE 174 AA; 18262 MW; 08B355A9A6F7F67A CRC64;

Query Match 5.0%; Score 9; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDKP 163
Db 150 GGDNYSDKP 158
|||||
|

RESULT 15
SOD1_SALTY STANDARD; PRT; 177 AA.
AC P53636; O33803; O50545;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Superoxide dismutase [Cu-Zn] 1 precursor (EC 1.15.1.1) (sodC1).
GN Names=sodC1; Synonyms=sodC; OrderedLocusNames=STM1044;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=ST4/74;
RX MEDLINE=98025474; PubMed=9379906;
RA Farrant J.L., Sansone A., Canvin J.R., Pallen M.J., Langford P.R.,
RA Wallis T.S., Dougan G., Kroll J.S.;

```

```

RT "Bacterial copper- and zinc-cofactored superoxide dismutase
RL contributes to the pathogenesis of systemic salmonellosis.";
RC Mol. Microbiol. 25:785-796(1997).
CC [2]
CC SEQUENCE FROM N.A.
RX MEDLINE=98054349; PubMed=9391141;
RA De Groote M.A., Ochsen U.A., Shiloh M.U., Nathan C., McCord J.M.,
RA Dinauer M.C., Libby S.J., Vazquez-Torres A., Xu Y., Fang F.C.;
RA "Periplasmic superoxide dismutase protects Salmonella from products of
RT phagocyte NADPH-oxidase and nitric oxide synthase.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:13997-14001(1997).
CC [3]
CC SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC [4]
CC SEQUENCE OF 77-163 FROM N.A.
RX STRAIN=ST4/74;
CC MEDLINE=97023146; PubMed=8869506;
RA Canvin J., Langford P.R., Wilks K.E., Kroll J.S.;
RT "Identification of sodC encoding periplasmic [Cu,Zn]-superoxide
RL dismutase in Salmonella.";
CC FEMS Microbiol. Lett. 136:215-220(1996).
CC [5]
CC X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS)
RX MEDLINE=20428907; PubMed=10370746; DOI=10.1006/jmbi.2000.4074;
RA Pesce A., Battistoni A., Trojano M.E., Polizio F., Nardini M.,
RA Kroll J.S., Langford P.R., O'Neill P., Sette M., Desideri A.,
RA Bolognesi M.;
RT "Functional and crystallographic characterization of Salmonella
RL Typhimurium Cu,Zn superoxide dismutase coded by the sodC1 virulence
RT gene.";
RL J. Mol. Biol. 302:465-478(2000).
CC -|- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -|- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit.
CC -|- SUBUNIT: Monomer (By similarity).
CC -|- SUBCELLULAR LOCATION: Periplasmic.
CC -|- MISCELLANEOUS: Encoded by a cryptic bacteriophage.
CC -|- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y13121; CAA73588.1; -.
CC FIRM; AF007380; AAB62385.1; -.
CC FIRM; AE008743; AAL19978.1; -.
CC EMBL; X94327; CAA63988.1; -.
CC PDB; 1EQW; X-ray; A/B/C/D=22-177.
CC StyGene; SGL0705; sodC1.
CC InterPro; IPR001424; SOD_CU_ZN.
CC ProDom; PD000469; SOD_CU_ZN; 1.
CC PROSITE; PS00087; SOD_CU_ZN_1; FALSE_NEG.
CC PROSITE; PS00332; SOD_CU_ZN_2; 1.
CC 3D-structure; Antioxidant; Complete proteome; Copper; Metal-binding;
KW Oxidoreductase; Periplasmic; Signal; Zinc.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 177 Superoxide dismutase [Cu-Zn] 1.

```

```

FT METAL 69 69 Copper.
FT METAL 71 71 Copper.
FT METAL 94 94 Copper and zinc.
FT METAL 103 103 Zinc.
FT METAL 112 112 Zinc.
FT METAL 115 115 Zinc.
FT METAL 150 150 Copper.
FT DISULFID 76 172
FT CONFLICT 49 49
FT CONFLICT 148 148
FT STRAND 23 31
FT STRAND 36 48
FT TURN 49 50
FT STRAND 51 58
FT STRAND 64 66
FT STRAND 68 71
FT STRAND 79 80
FT STRAND 85 86
FT TURN 87 88
FT HELIX 89 91
FT STRAND 94 94
FT TURN 97 98
FT TURN 106 107
FT TURN 109 110
FT TURN 113 114
FT STRAND 115 115
FT STRAND 119 121
FT TURN 123 124
FT STRAND 127 127
FT STRAND 131 133
FT TURN 134 135
FT HELIX 139 142
FT TURN 143 144
FT STRAND 146 150
FT HELIX 162 165
FT STRAND 169 174
SQ SEQUENCE 177 AA; 18370 MW; 1EC743EE2AB38CAE CRC64;

Query Match 5.0%; Score 9; DB 1; Length 177;
Best Local Similarity 100.0%; Pred.No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDKP 163
Db 152 GGDNYSDKP 160

RESULT 16
Q704S7 PRELIMINARY; PRT; 177 AA.
AC Q704S7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Superoxide dismutase precursor (EC 1.15.1.1).
GN Name=sodCl;
OS Salmonella enterica subsp. enterica serovar Choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=119912;
RN [1]
RP SEQUENCE FROM N.A.
RA Ammendola S., Ajello M., Pasquali P., Kroll J.S., Langford P.R.,
RA Rotilio G., Valenti P., Battistoni A.;
RT "Differential contribution of sodCl and sodC2 to intracellular
RT survival and pathogenicity of Salmonella enterica subsp. enterica
RT serovar Choleraesuis.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620903; CAF06531.1; -.
KW Oxidoreductase; Signal.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 177 superoxide dismutase [Cu-Zn].
SQ SEQUENCE 177 AA; 18370 MW; 1EC743EE2AB38CAE CRC64;

Query Match 5.0%; Score 9; DB 2; Length 177;
Best Local Similarity 100.0%; Pred.No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDKP 163
Db 152 GGDNYSDKP 160

RESULT 17
CAF06531 PRELIMINARY; PRT; 177 AA.
AC CAF06531;
DT 10-MAY-2004 (TREMBlrel. 27, Created)
DT 10-MAY-2004 (TREMBlrel. 27, Last sequence update)
DT 10-MAY-2004 (TREMBlrel. 27, Last annotation update)
DE Superoxide dismutase precursor (EC 1.15.1.1).
GN SODCl.
OS Salmonella enterica subsp. enterica serovar Choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella; Salmonella enterica.
OX NCBI_TaxID=119912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AS0;
RA Ammendola S., Ajello M., Pasquali P., Kroll J.S., Langford P.R.,
RA Rotilio G., Valenti P., Battistoni A.;
RT "Differential contribution of sodCl and sodC2 to intracellular
RT survival and pathogenicity of Salmonella enterica subsp. enterica
RT serovar Choleraesuis.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620903; CAF06531.1; -.
KW Oxidoreductase; Signal.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 177 superoxide dismutase [Cu-Zn].
SQ SEQUENCE 177 AA; 18370 MW; 1EC743EE2AB38CAE CRC64;

Query Match 5.0%; Score 9; DB 2; Length 177;
Best Local Similarity 100.0%; Pred.No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GGDNYSDKP 163
Db 152 GGDNYSDKP 160

RESULT 18
Q6MR06 PRELIMINARY; PRT; 189 AA.
AC Q6MR06;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE SodC protein precursor (EC 1.15.1.1).
GN Name=sodC; OrderedLocusNames=Bd0295;
OS Bdellovibrio bacteriovorus
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164;

```

RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,  
 RA Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,  
 RA Sockett R.E., Schuster S.C.;  
 RT "A predator unmasked: life cycle of *Bdellovibrio bacteriovorus* from a  
 RL genomic perspective."; Science 303:689-692(2004).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 CC cells and which are toxic to biological systems (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
 DR EMBL; BX842646; CAE77952.1; -.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0004784; F:superoxide dismutase activity; IEA.  
 DR InterPro; IPR001424; SOD\_CU\_ZN.  
 DR Pfam; PF00080; SOD\_Cu; 1.  
 DR PRINTS; PR00068; CUZNDISMTASE.  
 DR ProDom; PD000469; SOD\_CU\_ZN; 1.  
 DR PROSITE; PS00087; SOD\_CU\_ZN\_1; 1.  
 DR PROSITE; PS00332; SOD\_CU\_ZN\_2; 1.  
 KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Signal;  
 KW Zinc.  
 FT SIGNAL 1 14 Potential.  
 SQ SEQUENCE 189 AA; 19587 MW; CEEFF88AA6D3AD82 CRC64;  
 Query Match 5.0%; Score 9; DB 2; Length 189;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 GARIACGVI 178  
 Db |||||  
 178 GARIACGVI 186

RESULT 19  
 CAE77952 PRELIMINARY; PRT; 189 AA.  
 ID CAE77952;  
 AC CAE77952;  
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE SODc protein precursor (EC 1.15.1.1).  
 GN SODC OR BD0295.  
 OS *Bdellovibrio bacteriovorus*.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;  
 OC Bellowibrionaceae; Bdellovibrio.  
 OX NCBI\_TaxID=959;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;  
 RX PubMed=14752164;  
 RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,  
 RA Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,  
 RA Sockett R.E., Schuster S.C.;  
 RT "A predator unmasked: life cycle of *Bdellovibrio bacteriovorus* from a  
 RL genomic perspective."; Science 303:689-692(2004).  
 CC EMBL; BX842646; CAE77952.1; -.  
 CC GO; GO:0016491; F:oxidoreductase activity; IEA.  
 CC GO; GO:0004784; F:superoxide dismutase activity; IEA.  
 CC InterPro; IPR001424; SOD\_CU\_ZN.  
 CC Pfam; PF00080; SOD\_Cu; 1.  
 CC PRINTS; PR00068; CUZNDISMTASE.  
 CC ProDom; PD000469; SOD\_CU\_ZN; 1.  
 CC PROSITE; PS00087; SOD\_CU\_ZN\_1; 1.  
 CC PROSITE; PS00332; SOD\_CU\_ZN\_2; 1.  
 KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Signal;  
 KW Zinc.  
 FT SIGNAL 1 14 Potential.  
 SQ SEQUENCE 189 AA; 19587 MW; CEEFF88AA6D3AD82 CRC64;

Query Match 5.0%; Score 9; DB 2; Length 189;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 GARIACGVI 178  
 Db |||||  
 178 GARIACGVI 186

RESULT 20

SODC ACTAC STANDARD; PRT; 87 AA.  
 ID SODC ACTAC  
 AC Q508L;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Fragment).  
 GN Names-sodc;  
 OS *Actinobacillus actinomycetemcomitans* (Haemophilus  
 OC actinomycetemcomitans).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Actinobacillus.  
 OX NCBI\_TaxID=714;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Y4;  
 RX MEDLINE=96118708; PubMed=7496539;  
 RA Kroll J.S., Langford P.R., Wilks K.E., Keil A.D.;  
 RT "Bacterial [Cu,Zn]-superoxide dismutase: phylogenetically distinct  
 RT from the eukaryotic enzyme, and not so rare after all!";  
 RL Microbiology 141:2271-2279(1995).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 CC cells and which are toxic to biological systems (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
 CC similarity).  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X81122; CAA58203.1; -.  
 DR PIR; I39485; I39485.  
 DR HSSP; P24702; 2APS.  
 DR InterPro; IPR001424; SOD\_CU\_ZN.  
 DR Pfam; PF00080; SOD\_Cu; 1.  
 DR ProDom; PD000469; SOD\_CU\_ZN; 1.  
 DR PROSITE; PS00087; SOD\_CU\_ZN\_1; PARTIAL.  
 DR PROSITE; PS00332; SOD\_CU\_ZN\_2; PARTIAL.  
 KW Antioxidant; Copper; Metal-binding; Oxidoreductase; Periplasmic; Zinc.  
 FT NON TER 1 1  
 FT METAL 18 18 Copper and zinc (By similarity).  
 FT METAL 27 27 Zinc (By similarity).  
 FT METAL 36 36 Zinc (By similarity).  
 FT METAL 39 39 Zinc (By similarity).  
 FT METAL 74 74 Copper (By similarity).  
 FT NON TER 87 87  
 SQ SEQUENCE 87 AA; 9287 MW; 6704173D25976550 CRC64;  
 Query Match 4.4%; Score 8; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGGDN 158  
 Db |||||  
 72 MIHAGGDN 79

RESULT 21  
 Q8PDZ3 PRELIMINARY; PRT; 165 AA.  
 ID Q8PDZ3;  
 AC Q8PDZ3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Superoxide dismutase.

```
GN Name=sodC2; OrderedLocusNames=XCC0191;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergro F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COPACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AE012115; AAM39510.1; -.
DR HSSP; P00441; 10ZU.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; SOD_Cu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_ZN_1; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 165 AA; 16625 MW; 0286D15ACE479A36 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 BHGFHIE 77
Db 50 BHGFHIE 57

RESULT 22
QBQW1
ID Q8PQW1 PRELIMINARY; PRT; 165 AA.
AC Q8PQW1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Superoxide dismutase.
GN Name=sodC2; OrderedLocusNames=XAC0210;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
```

```
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergro F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COPACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AE011646; AAM35102.1; -.
DR HSSP; P00441; 10ZU.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; SOD_Cu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_ZN_1; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 165 AA; 16678 MW; 28C6FB3DF09F7E81 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 BHGFHIE 77
Db 50 BHGFHIE 57

RESULT 23
QB7G06
ID Q87G06 PRELIMINARY; PRT; 171 AA.
AC Q87G06;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Superoxide dismutase, Cu-Zn.
GN OrderedLocusNames=VPA1514;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Naitima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RA "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
```



```

DR EMBL; AE005387; AAG56635.1; -.
DR EMBL; AP002558; BAB35778.1; -.
DR EMBL; X97766; CAAG6363.1; -.
DR PIR; C90923; C90923.
DR PIR; G85771; G85771.
DR PIR; JC6004; JC6004.
DR PDB; 1ESO; X-ray; @=20-173.
DR EchoBASE; EB3195; -.
DR EcoGene; EG13419; sodC.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; FALSE NEG.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW 3D-structure; Antioxidant; Complete proteome; Copper;
KW Direct protein sequencing; Metal-binding; Oxidoreductase; Periplasmic;
KW Signal; Zinc.
FT SIGNAL 1 19 Superoxide dismutase [Cu-Zn].
FT CHAIN 20 173 Copper.
FT METAL 67 67 Copper.
FT METAL 69 69 Copper.
FT METAL 92 92 Copper and zinc.
FT METAL 101 101 Zinc.
FT METAL 109 109 Zinc.
FT METAL 112 112 Zinc.
FT METAL 147 147 Copper.
FT DISULFID 74 169
FT STRAND 21 30
FT TURN 31 32
FT STRAND 33 46
FT TURN 47 48
FT STRAND 49 56
FT STRAND 62 64
FT STRAND 66 69
FT STRAND 77 77
FT STRAND 84 84
FT TURN 85 86
FT HELIX 87 89
FT STRAND 92 92
FT TURN 95 96
FT TURN 104 105
FT TURN 110 111
FT STRAND 112 112
FT STRAND 116 118
FT TURN 120 121
FT STRAND 124 124
FT STRAND 128 130
FT TURN 131 132
FT HELIX 136 139
FT TURN 140 141
FT STRAND 142 147
FT HELIX 159 162
FT STRAND 166 173
SQ SEQUENCE 173 AA; 17681 MW; 9A0CB65F03AAB197 CRC64;

Query Match 4.4%; Score 8; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 GEHGFHIH 76
Db 62 GEHGFHIH 69

RESULT 25
SODC_POHLE
ID SODC_POHLE STANDARD; PRT; 173 AA.
AC P00446;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Superoxide dismutase [Cu-Zn] precursor [EC 1.15.1.1].
GN Name=sodC;

```

```

OS Photobacterium leiognathi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=658;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87109348; PubMed=3805055;
RA Steinman H.M.;
RT "Bacteriocuprein superoxide dismutase of Photobacterium leiognathi.
RT Isolation and sequence of the gene and evidence for a precursor
RT form.";
RL J. Biol. Chem. 262:1882-1887(1987).
RN [2]
RP SEQUENCE OF 23-173.
RX MEDLINE=83289129; PubMed=6884993;
RA Steffens G.J., Bannister J.V., Bannister W.H., Flohe L., Gunzler W.A.,
RA Kim S.-W.A., Oetting F.;
RT "The primary structure of Cu-Zn superoxide dismutase from
RT Photobacterium leiognathi: evidence for a separate evolution of Cu-Zn
RT superoxide dismutase in bacteria.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:675-690(1983).
RN [3]
RP DISCUSSION OF POSSIBLE GENE TRANSFER FROM EUKARYOTES.
RX MEDLINE=85113139; PubMed=3855538;
RA Bannister J.V., Parker M.W.;
RT "The presence of a copper/zinc superoxide dismutase in the bacterium
RT Photobacterium leiognathi: a likely case of gene transfer from
RT eukaryotes to prokaryotes.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:149-152(1985).
RN [4]
RP DISCUSSION OF POSSIBLE GENE TRANSFER FROM EUKARYOTES.
RX Leunissen J.A.M., de Jong W.W.;
RT "Copper/zinc superoxide dismutase: how likely is gene transfer from
RT ponyfish to Photobacterium leiognathi.";
RL J. Mol. Evol. 23:250-258(1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97075068; PubMed=8917495;
RA Bourne Y., Redford S.M., Steinman H.M., Lepock J.R., Tainer J.A.,
RA Getzoff E.D.;
RT "Novel dimeric interface and electrostatic recognition in bacterial
RT Cu,Zn superoxide dismutase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12774-12779(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99096923; PubMed=9878406;
RA Bordo D., Matak D., Djinic-Carugo X., Rosano C., Pesce A.,
RA Bolognesi M., Stroppolo M.E., Falconi M., Battistoni A., Desideri A.;
RT "Evolutionary constraints for dimer formation in prokaryotic Cu,Zn
RT superoxide dismutase.";
RL J. Mol. Biol. 285:283-296(1999).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02658; AAA25632.1; -.
DR PIR; A26689; DSFOCL.
DR PDB; 1BZO; X-ray; A=23-173.
DR PDB; 1IB5; X-ray; A=23-173.
DR PDB; 1IBB; X-ray; A=23-173.
DR PDB; 1IBD; X-ray; A=23-173.

```

```

DR PDB; 1IBF; X-ray; A=23-173.
DR PDB; 1IBH; X-ray; A=23-173.
DR PDB; 1OAJ; X-ray; A=23-173.
DR PDB; 1OAJ; X-ray; A=23-173.
DR PDB; 1YAI; X-ray; A/B/C=23-173.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; SOD_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW 3D-structure; Antioxidant; Copper; Direct protein sequencing;
KW Metal-binding; Oxidoreductase; Periplasmic; Signal; Zinc.
FT SIGNAL 1 22
FT CHAIN 23 173 Superoxide dismutase [Cu-Zn].
FT METAL 67 67 Copper.
FT METAL 69 69 Copper.
FT METAL 92 92 Copper and zinc.
FT METAL 101 101 Zinc.
FT METAL 110 110 Zinc.
FT METAL 113 113 Zinc.
FT METAL 147 147 Copper.
FT DISULFID 74 169
FT STRAND 24 31
FT TURN 32 34
FT STRAND 37 46
FT TURN 47 48
FT STRAND 49 56
FT STRAND 62 64
FT STRAND 66 69
FT STRAND 77 79
FT TURN 80 81
FT STRAND 82 84
FT TURN 85 86
FT HELIX 87 89
FT STRAND 92 92
FT TURN 95 96
FT TURN 104 105
FT TURN 111 112
FT STRAND 113 113
FT STRAND 117 119
FT TURN 121 122
FT STRAND 125 125
FT STRAND 129 131
FT TURN 132 133
FT TURN 136 139
FT TURN 140 141
FT STRAND 143 147
FT HELIX 159 162
FT STRAND 166 171
SQ SEQUENCE 173 AA; 18109 MW; 5931576F1E2A8F47 CRC64;

Query Match 4.4%; Score 8; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGSDN 158
Db 145 MIHAGSDN 152
|||||

RESULT 26
SODC_SALTY STANDARD; PRT; 173 AA.
AC O68901;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Superoxide dismutase [Cu-Zn] 2 precursor (EC 1.15.1.1) (sodCII).
GN Name=sodC; Synonyms=sodC2; OrderedLocusNames=STM1440;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OC NCBI_TaxID=602;

[1]
SEQUENCE FROM N.A.
RN STRAIN=ATCC 14028S;
RX MEDLINE=99307439; PubMed=10377444;
RA Fang F.C., Degroote M.A., Foster J.W., Bauml A.J., Ochsner U.,
RA Testerman T., Bearson S., Giarl J.C., Xu Y., Campbell G., Laessig T.;
RT "Virulent Salmonella typhimurium has two periplasmic Cu, Zn-superoxide
dismutases.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:7502-7507(1999).
[2]
SEQUENCE FROM N.A.
RN STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
(Potential).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF056931; RAC13559.1; -.
CC EMBL; AE008762; AAL20362.1; -.
CC HSPB; P53635; IESO.
CC StyGene; SG77777; sodC.
CC InterPro; IPR001424; SOD_CU_ZN.
CC Pfam; PF00080; SOD_Cu; 1.
CC ProDom; PD000469; SOD_CU_ZN; 1.
CC PROSITE; PS00087; SOD_CU_ZN_1; FALSE_NEG.
CC PROSITE; PS00332; SOD_CU_ZN_2; 1.
CC Antioxidant; Complete proteome; Copper; Metal-binding; Oxidoreductase;
KW Periplasmic; Signal; Zinc.
FT SIGNAL 1 19 By similarity.
FT CHAIN 20 173 Superoxide dismutase [Cu-Zn] 2.
FT METAL 67 67 Copper (By similarity).
FT METAL 69 69 Copper (By similarity).
FT METAL 92 92 Copper and zinc (By similarity).
FT METAL 101 101 Zinc (By similarity).
FT METAL 109 109 Zinc (By similarity).
FT METAL 112 112 Zinc (By similarity).
FT METAL 147 147 Copper (By similarity).
FT DISULFID 74 169 By similarity.
SQ SEQUENCE 173 AA; 17737 MW; 5FDC9F6F6B3EF CRC64;

Query Match 4.4%; Score 8; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GERGFHIIH 76
Db 62 GERGFHIIH 69
|||||

RESULT 27
Q751L8

```



```

ID Q751L8 PRELIMINARY; PRT; 173 AA.
AC Q751L8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGJ321Wp.
GN Name=AGL321W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RX PubMed=15001715;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AE016905; AAS54170.1; -.
DR AGD; AGL321W; -.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 173 AA; 18116 MW; A3FF0E9F564E9CBA CRC64;

Query Match 4.4%; Score 8; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 EHGPHIHE 77
Db 62 EHGPHIHE 69

RESULT 28
Q704S6 PRELIMINARY; PRT; 173 AA.
ID Q704S6;
AC Q704S6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGJ321Wp.
GN Name=sodC2;
OS Salmonella enterica subsp. enterica serovar Choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=119912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AL627271; CAD01927.1; -.
DR EMBL; AE016838; AAO68958.1; -.
DR HSP; P53635; IESO.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00332; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.

```

```

DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Copper; Metal-binding; Oxidoreductase; Signal; Zinc.
FT SIGNAL 19 Potential.
FT CHAIN 20 173 superoxide dismutase [Cu-Zn].
SQ SEQUENCE 173 AA; 17721 MW; 511A19E8057F1255 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 GEHGFHIIH 76
Db 62 GEHGFHIIH 69

RESULT 29
Q8Z6P6 PRELIMINARY; PRT; 173 AA.
ID Q8Z6P6; Q7CA25;
AC Q8Z6P6; Q7CA25;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Copper-zinc superoxide dismutase.
GN Name=sodC; OrderedLocusNames=STY1682, t1308;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AL627271; CAD01927.1; -.
DR EMBL; AE016838; AAO68958.1; -.
DR HSP; P53635; IESO.
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00332; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.

```



```
SQ SEQUENCE 173 AA; 17735 MW; 00A5A8B6AF25B4EF CRC64;
Query Match 4.4%; Score 8; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GEHGFHIIH 76
DB 62 GEHGFHIIH 69

RESULT 30
ID AAS54170 PRELIMINARY; PRT; 173 AA.
AC AAS54170;
DT 23-APR-2004 (TrEMBLrel. 27, Created)
DT 23-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 23-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE AGL321WP.
GN AGL321WP.
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RX PubMed=15001715;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavler A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AF016905; AAS54170.1; -.
SQ SEQUENCE 173 AA; 18116 MW; A3FF0E9F564E9CBA CRC64;

Query Match 4.4%; Score 8; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHGFIHIE 77
DB 62 EHGFIHIE 69

RESULT 31
ID CAF06532 PRELIMINARY; PRT; 173 AA.
AC CAF06532;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Superoxide dismutase precursor (EC 1.15.1.1).
GN SODC2.
OS Salmonella enterica subsp. enterica serovar Choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella; Salmonella enterica.
OX NCBI_TaxID=119912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A50;
RA Ammendola S., Ajello M., Pasquali P., Kroll J.S., Langford P.R.,
RA Rotilio G., Valenti P., Battistoni A.;
RT "Differential contribution of sodC1 and sodC2 to intracellular
RT survival and pathogenicity of Salmonella enterica subsp. enterica
RT serovar Choleraesuis.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620904; CAF06532.1; -.
KW Oxidoreductase; Signal.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 173 superoxide dismutase [Cu-Zn].
SQ SEQUENCE 173 AA; 17721 MW; 511A19E8057F1255 CRC64;
```

```
Query Match 4.4%; Score 8; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GEHGFHIIH 76
DB 62 GEHGFHIIH 69

RESULT 32
ID Q7AEAI PRELIMINARY; PRT; 175 AA.
AC Q7AEAI;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative copper/zinc-superoxide dismutase.
GN OrderedLocusNames=ECs1989;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohkubo E., Nakayama K., Murata T., Tanaka M., Tobé T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AP002556; BAB35412.1; -.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; SOD_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 175 AA; 18275 MW; 1098B6A5F7B9478E CRC64;

Query Match 4.4%; Score 8; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 PLGGGGAR 172
DB 161 PLGGGGAR 168

RESULT 33
ID Q7AFX5 PRELIMINARY; PRT; 175 AA.
AC Q7AFX5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative copper/zinc-superoxide dismutase.
GN OrderedLocusNames=ECs1120;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
```

REX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Okabayashi K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
RA Kihara S., Shiba T., Hattori M., Shinagawa H.,  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
CC -!- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems (By similarity).  
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
CC EMBL; AP002554; BAB34543.1; -.  
DR InterPro; IPR001424; SOD\_CU\_ZN.  
DR Pfam; PF00080; Sod\_Cu; 1. SOD\_CU\_ZN; 1.  
DR ProDom; PD000469; SOD\_CU\_ZN; 1.  
DR PROSITE; PS00332; SOD\_CU\_ZN 2; 1.  
KW Copper; Metal-binding; Oxidoreductase; Zinc.  
SQ SEQUENCE 175 AA; 18259 MW; 00D8A6F5B7B94792 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 165 PLGGGGAR 172  
|||||

Db 161 PLGGGGAR 168  
|||||

## RESULT 34

ID Q7UCE0 PRELIMINARY; PRT; 178 AA.  
AC Q7UCE0;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Superoxide dismutase (Cu-Zn).  
GN Name=sodC; OrderedLocusNames=S1805;  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Shigella.  
OX NCBI\_TaxID=623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2457T;  
RX MEDLINE=22590274; PubMed=12704152;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
RA Schwartz D.C., Blattner F.R.;  
RT "Complete genome sequence and comparative genomics of Shigella  
RT flexneri serotype 2a strain 2457T.";  
RL Infect. Immun. 71:2775-2786(2003).  
CC -!- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems (By similarity).  
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
CC EMBL; AE016983; AAP17140.1; -.  
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.  
DR GO; GO:0046872; P:metal ion binding; IEA.  
DR GO; GO:0006801; P:superoxide metabolism; IEA.  
DR InterPro; IPR001424; SOD\_CU\_ZN.  
DR Pfam; PF00080; Sod\_Cu; 1.  
DR ProDom; PD000469; SOD\_CU\_ZN; 1.  
DR PROSITE; PS00332; SOD\_CU\_ZN 2; 1.  
KW Copper; Metal-binding; Oxidoreductase; Zinc.  
SQ SEQUENCE 178 AA; 18167 MW; DC43B64712E8FA65 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 178;

Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 69 GEHGFHLH 76  
|||||  
Db 67 GEHGFHLH 74  
|||||  
RESULT 35  
SODC\_HAEIN  
ID SODC\_HAEIN STANDARD; PRT; 187 AA.  
AC P25841;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Superoxide dismutase [Cu-Zn] like precursor.  
DE Name=sodC;  
OS Haemophilus influenzae;  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCTC 8468 / Serotype B;  
RX MEDLINE=92041655; PubMed=1938942;  
RA Kroll J.S., Langford P.R., Loynds B.M.;  
RT "Copper-zinc superoxide dismutase of Haemophilus influenzae and H.  
RT parainfluenzae.";  
RL J. Bacteriol. 173:7449-7457(1991).  
CC -!- FUNCTION: This protein lacks enzymatic activity (probably because  
CC of the presence of a tyrosine instead of a histidine at residue  
CC 82).  
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
CC similarity).  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).  
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; M84012; AAA24953.1; -.  
DR PIR; A41654; A41654.  
DR HSRP; P24702; 2APS.  
DR InterPro; IPR001424; SOD\_CU\_ZN.  
DR Pfam; PF00080; Sod\_Cu; 1.  
DR ProDom; PD000469; SOD\_CU\_ZN; 1. FALSE\_NEG.  
DR PROSITE; PS00087; SOD\_CU\_ZN 1; FALSE\_NEG.  
DR PROSITE; PS00332; SOD\_CU\_ZN 2; 1.  
KW Copper; Metal-binding; Periplasmic; Signal; Zinc.  
FT SIGNAL 1 23 Potential.  
FT CHAIN 24 187 Superoxide dismutase [Cu-Zn] like.  
FT METAL 80 80 Copper (By similarity).  
FT SITE 82 82 ANCESTRAL COPPER-BINDING SITE.  
FT METAL 105 105 Copper and zinc (By similarity).  
FT METAL 114 114 Zinc (By similarity).  
FT METAL 123 123 Zinc (By similarity).  
FT METAL 126 126 Zinc (By similarity).  
FT METAL 161 161 Copper (By similarity).  
FT DISULFID 87 183 By similarity.  
SQ SEQUENCE 187 AA; 19536 MW; 3EE95EFFD52425B3 CRC64;

Query Match 4.4%; Score 8; DB 1; Length 187;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 MIHAGDGN 158  
|||||

Db 159 MIHAGSDN 166

```
RESULT 36
SODC_HAEP
ID SODC_HAEP STANDARD; PRT; 187 AA.
AC P25842;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1).
GN Name=sodC;
OS Haemophilus parainfluenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=729;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1391;
RX MEDLINE=92041655; PubMed=1938942;
RA Kroll J.S., Langford P.R., Loynds B.M.;
RT "Copper-zinc superoxide dismutase of Haemophilus influenzae and H.
RT parainfluenzae.";
RL J. Bacteriol. 173:7449-7457(1991).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- FUNCTION: May confer survival advantage by accelerating
CC dismutation of superoxide of environmental origin to hydrogen
CC peroxide, disruptive to the normal mucociliary clearance process
CC in the host.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
-----
DR EMBL; M84013; AAA24954.1; -.
DR PIR; B41654; B41654.
DR HSSP; P24702; 2APS.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Antioxidant; Copper; Metal-binding; Oxidoreductase; Periplasmic;
KW Signal; Zinc.
FT SIGNAL 1 23 Potential.
FT CHAIN 24 187 Superoxide dismutase [Cu-Zn].
FT METAL 80 80 Copper (By similarity).
FT METAL 82 82 Copper (By similarity).
FT METAL 105 105 Copper and zinc (By similarity).
FT METAL 114 114 Zinc (By similarity).
FT METAL 123 123 Zinc (By similarity).
FT METAL 126 126 Zinc (By similarity).
FT METAL 161 161 Copper (By similarity).
FT DISULFID 87 183 By similarity.
SQ SEQUENCE 187 AA; 19510 MW; A0C3A61EFAF201D5 CRC64;

Query Match 4.4%; Score 8; DB 1; Length 187;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGSDN 158
|||||
```

Db 159 MIHAGSDN 166

```
RESULT 37
Q714V4
ID Q714V4 PRELIMINARY; PRT; 187 AA.
AC Q714V4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SodC.
GN Name=sodC;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=700222;
RX MEDLINE=22999298; PubMed=14638817;
RA Satola S.W., Schirmer P.L., Farley M.M.;
RT "Genetic analysis of the capsule locus of Haemophilus influenzae
RT serotype f.";
RL Infect. Immun. 71:7202-7207(2003).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR EMBL; AF549211; AAQ12654.1; -.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; UNKNOWN_1.
KW Copper; Metal-binding; Oxidoreductase; Zinc.
SQ SEQUENCE 187 AA; 19696 MW; 88E2040A031AF8F7 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 MIHAGSDN 158
|||||
```

RESULT 38

```
AAQ12654
ID AAQ12654 PRELIMINARY; PRT; 187 AA.
AC AAQ12654;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE SodC.
GN SODC.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=700222;
RX MEDLINE=22999298; PubMed=14638817;
RA Satola S.W., Schirmer P.L., Farley M.M.;
RT "Genetic analysis of the capsule locus of Haemophilus influenzae
RT serotype f.";
RL Infect. Immun. 71:7202-7207(2003).
DR EMBL; AF549211; AAQ12654.1; -.
SQ SEQUENCE 187 AA; 19696 MW; 88E2040A031AF8F7 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 18;
```

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 MIHAGGN 158  
|||||  
Db 159 MIHAGGN 166

## RESULT 39

Q83RB4 PRELIMINARY; PRT; 190 AA.  
AC Q83RB4  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Superoxide dismutase (Cu-Zn).  
GN Name=sodC; OrderedLocusNames=SF1673;  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Shigella.  
OX NCBI\_TaxID=623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=301 / Serotype 2a;  
RA MEDLINE=22272406; PubMed=12384590;  
RX Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.;  
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
RT through comparison with genomes of Escherichia coli K12 and O157.";  
RL Nucleic Acids Res. 30:4432-4441(2002).  
CC -!- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems (By similarity).  
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -!- COPACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
DR EMBL; AE015188; AAN43254.1; -.  
DR HSSP; P53635; 1ES0.  
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.  
DR GO; GO:0046872; F:metal ion binding; IEA.  
DR GO; GO:0006801; P:superoxide metabolism; IEA.  
DR InterPro; IPR001424; SOD\_CU\_ZN.  
DR Pfam; PF00080; Sod\_Cu; 1.  
DR ProDom; PD000469; SOD\_CU\_ZN; 1.  
DR PROSITE; PS00332; SOD\_CU\_ZN\_2; 1.  
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.  
SQ SEQUENCE 190 AA; 19661 MW; 6217FD0EE596E253 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred.No.18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 GEHGFHIIH 76  
|||||  
Db 79 GEHGFHIIH 86

## RESULT 40

Q8FH80 PRELIMINARY; PRT; 190 AA.  
AC Q8FH80  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Superoxide dismutase (EC 1.15.1.1).  
GN Name=sodC; OrderedLocusNames=c2038;  
OS Escherichia coli O6.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=217992;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=06:H1 / CFT073 / ATCC 700928 / UPEC;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.F.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
RT of uropathogenic Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
CC -!- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems (By similarity).  
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -!- COPACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.  
DR EMBL; AE016761; AAN80498.1; -.  
DR HSSP; P53635; 1ES0.  
DR GO; GO:0004785; F:copper, zinc superoxide dismutase activity; IEA.  
DR GO; GO:0046872; F:metal ion binding; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006801; P:superoxide metabolism; IEA.  
DR InterPro; IPR001424; SOD\_CU\_ZN.  
DR Pfam; PF00080; Sod\_Cu; 1.  
DR ProDom; PD000469; SOD\_CU\_ZN; 1.  
DR PROSITE; PS00332; SOD\_CU\_ZN\_2; 1.  
KW Complete proteome; Copper; Metal-binding; Oxidoreductase; Zinc.  
SQ SEQUENCE 190 AA; 19617 MW; 6217F2DAAE1A8E23 CRC64;

Query Match 4.4%; Score 8; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred.No.18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 GEHGFHIIH 76  
|||||  
Db 79 GEHGFHIIH 86

Search completed: October 26, 2004, 10:03:24  
Job time : 237 secs